

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2004, 17:26:21 ; Search time 20 Seconds  
(without alignments)  
1701.076 Million cell updates/sec

Title: US-09-802-285A-2

Perfect score: 3494  
Sequence: 1 MTTKIFRIIVFAVIALSSG.....KGKLNLTITNGKQOLVLP 659

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/2/iaa/5A COMB.pap:\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB.pap:\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB.pap:\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB.pap:\*  
5: /cgn2\_6/ptodata/2/iaa/6C COMB.pap:\*  
6: /cgn2\_6/ptodata/2/iaa/6D COMB.pap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3494	100.0	659	1 US-08-258-639A-4	Sequence 4, Appli
2	3494	100.0	659	2 US-08-900-951-4	Sequence 4, Appli
3	3494	100.0	659	5 PCT-US95-07391A-4	Sequence 4, Appli
4	164	4.7	649	4 US-09-134-000C-5302	Sequence 5302, Ap
5	126.5	3.6	698	4 US-09-134-001C-3632	Sequence 3632, Ap
6	124	3.5	23	1 US-08-258-639A-9	Sequence 9, Appli
7	124	3.5	23	2 US-08-900-951-9	Sequence 9, Appli
8	124	3.5	23	5 PCT-US95-07391A-9	Sequence 9, Appli
9	119	3.4	23	1 US-08-258-639A-10	Sequence 10, Appl
10	119	3.4	23	2 US-08-900-951-10	Sequence 10, Appl
11	119	3.4	23	5 PCT-US95-07391A-10	Sequence 10, Appl
12	119	3.4	1138	1 US-07-973-320-2	Sequence 2, Appli
13	119	3.4	1138	1 US-07-973-320-4	Sequence 2, Appli
14	118	3.4	772	1 US-08-258-639A-2	Sequence 2, Appli
15	118	3.4	772	2 US-08-900-951-2	Sequence 2, Appli
16	118	3.4	772	5 PCT-US95-07391A-2	Sequence 2, Appli
17	117.5	3.4	1028	4 US-09-543-681A-7181	Sequence 7181, Ap
18	116	3.3	4536	4 US-09-180-422B-27	Sequence 27, Appl
19	116	3.3	4536	4 US-09-079-030-1	Sequence 1, Appli
20	115.5	3.3	1168	1 US-08-620-717A-9	Sequence 9, Appli
21	114.5	3.3	1167	1 US-08-485-568A-6	Sequence 6, Appli
22	114.5	3.3	1167	2 US-08-590-554A-6	Sequence 6, Appli
23	114.5	3.3	1167	2 US-08-184-223-6	Sequence 6, Appli
24	113.5	3.2	1398	1 US-08-750-532-9	Sequence 9, Appli
25	113.5	3.2	1398	3 US-08-894-818B-8	Sequence 8, Appli
26	113.5	3.2	1398	4 US-09-445-472-6	Sequence 6, Appli
27	111	3.2	927	4 US-09-134-001C-4831	Sequence 4831, Ap

28 110 3.1 4563 4 US-09-108-006C-1 Sequence 1, Appli  
29 109 3.1 951 3 US-08-816-346-58 Sequence 58, Appli  
30 109 3.1 951 3 US-09-335-411-58 Sequence 58, Appli  
31 109 3.1 952 2 US-08-788-674-5 Sequence 5, Appli  
32 109 3.1 952 3 US-08-816-346-4 Sequence 4, Appli  
33 109 3.1 952 3 US-09-335-411-4 Sequence 4, Appli  
34 108.5 3.1 1096 4 US-09-134-000C-5764 Sequence 5764, Ap  
35 108 3.1 1015 4 US-09-134-000C-6204 Sequence 6204, Ap  
36 106 3.0 884 4 US-09-328-352-4598 Sequence 4598, Ap  
37 106 3.0 945 4 US-09-198-452A-1030 Sequence 1030, Ap  
38 106 3.0 1375 3 US-09-210-361-4 Sequence 4, Appli  
39 106 3.0 1375 4 US-09-740-274-4 Sequence 4, Appli  
40 105 3.0 443 4 US-09-328-352-6322 Sequence 6322, Ap  
41 105 3.0 741 4 US-09-252-991A-22440 Sequence 22440, A  
42 105 3.0 965 4 US-09-437-277-3 Sequence 3, Appli  
43 104 3.0 1742 4 US-09-386-962C-4 Sequence 4, Appli  
44 104 3.0 1849 4 US-08-851-567B-49 Sequence 49, Appl  
45 104 3.0 2516 4 US-08-851-567B-47 Sequence 47, Appl

#### ALIGNMENTS

RESULT 1  
US-08-258-639A-4  
; Sequence 4, Application US/08258639A  
; Patent No. 5681733  
; GENERAL INFORMATION:  
; APPLICANT: Su, Hongsheng  
; APPLICANT: Blain, Francoise  
; APPLICANT: Bennett, Clark  
; APPLICANT: Gu, Kangfu  
; APPLICANT: Zimmermann, Joseph  
; APPLICANT: Musil, Roy  
; TITLE OF INVENTION: Nucleic Acid Sequences And Expression  
; TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From  
; TITLE OF INVENTION: Flavobacterium heparinum  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hale and Dorr  
; STREET: 1455 Pennsylvania Avenue, N.W.  
; CITY: Washington, D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/08/258,639A  
; FILING DATE: 10 JUNE 1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Healey, William J.  
; REGISTRATION NUMBER: 36,160  
; REFERENCE/DOCKET NUMBER: 104385.116  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)942-8400  
; TELEFAX: (202)942-8484  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 659 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-258-639A-4

Query Match 100.0%; Score 3494; DB 1; Length 659;  
Best Local Similarity 100.0%; Pred. No. 4.1e-308;  
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTTKIFRIIVFAVIALSSGILQSSSITKDFHINLEYSGLKYNKVAAGNYDDAA 60

```
Db 1 MTTKIFKRIIVFAVIALSSGNILAQSSITRKDFDHNLEYSGLKKNKAAVAGNYDDAA 60
QY 61 KALLAYREKSKAREPDFSNAEKPADIROPIDKVTREMAKALVHQFQPHKGYGFDYDK 120
Db 61 KALLAYREKSKAREPDFSNAEKPADIROPIDKVTREMAKALVHQFQPHKGYGFDYDK 120
QY 121 DINQWMPVKDNEVRQLHRVQWQAMALVYHATGDEKYAREWVYQYSDWARKNPGLSQ 180
Db 121 DINQWMPVKDNEVRQLHRVQWQAMALVYHATGDEKYAREWVYQYSDWARKNPGLSQ 180
QY 181 DNDKFVRPLEVSDRVQSLPPTFSLFVNSPAFTPAFLMEFLNSYHQADYLSHYAEQGN 240
Db 181 DNDKFVRPLEVSDRVQSLPPTFSLFVNSPAFTPAFLMEFLNSYHQADYLSHYAEQGN 240
QY 241 HRLFEAQRNLFAGVSPPEFKDSPRWRQTGISVLNTEIKKQVYADGMQFELSPIYHVAAD 300
Db 241 HRLFEAQRNLFAGVSPPEFKDSPRWRQTGISVLNTEIKKQVYADGMQFELSPIYHVAAD 300
QY 301 IFLKAYGSAKRVNLEKEFFQSYVQTVENMIMALISISLDPYNTPMFGDSWITDKNFRMAQ 360
Db 301 IFLKAYGSAKRVNLEKEFFQSYVQTVENMIMALISISLDPYNTPMFGDSWITDKNFRMAQ 360
QY 361 FASWARVFPANQAIKYFATDGKQKAPNFKLSKALSNAAGFYTFRSKWDKKNATVMVLKASPP 420
Db 361 FASWARVFPANQAIKYFATDGKQKAPNFKLSKALSNAAGFYTFRSKWDKKNATVMVLKASPP 420
QY 421 GEFHAQPDNGTGFELFIKGRNFTPDAGVFVYSGDEAIMKLNWYRQTRIHSITLTDNQNMV 480
Db 421 GEFHAQPDNGTGFELFIKGRNFTPDAGVFVYSGDEAIMKLNWYRQTRIHSITLTDNQNMV 480
QY 481 ITKARQNKQWETGNLQVLTYNPSYPLNDHQRSVLFINKKYFLVIDRAIGEATNGLGVHW 540
Db 481 ITKARQNKQWETGNLQVLTYNPSYPLNDHQRSVLFINKKYFLVIDRAIGEATNGLGVHW 540
QY 541 QLKEDSNPVDKTKNRVYTYTYRDGNNLMIQSLNADRTSLNEEBGKSVYVYNKELKRPFAV 600
Db 541 QLKEDSNPVDKTKNRVYTYTYRDGNNLMIQSLNADRTSLNEEBGKSVYVYNKELKRPFAV 600
QY 601 FEKPKKNAGTQNFVSIVPYDGGQKAPEISIRENKGNDFEKGLNLTITINGKQOOLVLP 659
Db 601 FEKPKKNAGTQNFVSIVPYDGGQKAPEISIRENKGNDFEKGLNLTITINGKQOOLVLP 659
```

```
RESULT 2
US-08-900-951-4
; Sequence 4, Application US/08900951
; Patent No. 5919693
; GENERAL INFORMATION:
; APPLICANT: Su, Hongsheng
; APPLICANT: Blain, Francoise
; APPLICANT: Bennett, Clark
; APPLICANT: Gu, Kangfu
; APPLICANT: Zimmermann, Joseph
; APPLICANT: Musil, Roy
; TITLE OF INVENTION: Nucleic Acid Sequences And Expression
; TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From
; TITLE OF INVENTION: Flavobacterium heparinum
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC DOS/MS DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,951
; FILING DATE:
```

```
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/258,639
; FILING DATE: 10 JUNE 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Healey, William J.
; REGISTRATION NUMBER: 36,160
; REFERENCE/DOCKET NUMBER: 104385.116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)942 8400
; TELEFAX: (202)942 8484
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-900-951-4

Query Match 100.0%; Score 3494; DB 2; Length 659;
Best Local Similarity 100.0%; Pred. No. 4.1e-308;
Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTTKIFKRIIVFAVIALSSGNILAQSSITRKDFDHNLEYSGLKKNKAAVAGNYDDAA 60
Db 1 MTTKIFKRIIVFAVIALSSGNILAQSSITRKDFDHNLEYSGLKKNKAAVAGNYDDAA 60
QY 61 KALLAYREKSKAREPDFSNAEKPADIROPIDKVTREMAKALVHQFQPHKGYGFDYDK 120
Db 61 KALLAYREKSKAREPDFSNAEKPADIROPIDKVTREMAKALVHQFQPHKGYGFDYDK 120
QY 121 DINQWMPVKDNEVRQLHRVQWQAMALVYHATGDEKYAREWVYQYSDWARKNPGLSQ 180
Db 121 DINQWMPVKDNEVRQLHRVQWQAMALVYHATGDEKYAREWVYQYSDWARKNPGLSQ 180
QY 181 DNDKFVRPLEVSDRVQSLPPTFSLFVNSPAFTPAFLMEFLNSYHQADYLSHYAEQGN 240
Db 181 DNDKFVRPLEVSDRVQSLPPTFSLFVNSPAFTPAFLMEFLNSYHQADYLSHYAEQGN 240
QY 241 HRLFEAQRNLFAGVSPPEFKDSPRWRQTGISVLNTEIKKQVYADGMQFELSPIYHVAAD 300
Db 241 HRLFEAQRNLFAGVSPPEFKDSPRWRQTGISVLNTEIKKQVYADGMQFELSPIYHVAAD 300
QY 301 IFLKAYGSAKRVNLEKEFFQSYVQTVENMIMALISISLDPYNTPMFGDSWITDKNFRMAQ 360
Db 301 IFLKAYGSAKRVNLEKEFFQSYVQTVENMIMALISISLDPYNTPMFGDSWITDKNFRMAQ 360
QY 361 FASWARVFPANQAIKYFATDGKQKAPNFKLSKALSNAAGFYTFRSKWDKKNATVMVLKASPP 420
Db 361 FASWARVFPANQAIKYFATDGKQKAPNFKLSKALSNAAGFYTFRSKWDKKNATVMVLKASPP 420
QY 421 GEFHAQPDNGTGFELFIKGRNFTPDAGVFVYSGDEAIMKLNWYRQTRIHSITLTDNQNMV 480
Db 421 GEFHAQPDNGTGFELFIKGRNFTPDAGVFVYSGDEAIMKLNWYRQTRIHSITLTDNQNMV 480
QY 481 ITKARQNKQWETGNLQVLTYNPSYPLNDHQRSVLFINKKYFLVIDRAIGEATNGLGVHW 540
Db 481 ITKARQNKQWETGNLQVLTYNPSYPLNDHQRSVLFINKKYFLVIDRAIGEATNGLGVHW 540
QY 541 QLKEDSNPVDKTKNRVYTYTYRDGNNLMIQSLNADRTSLNEEBGKSVYVYNKELKRPFAV 600
Db 541 QLKEDSNPVDKTKNRVYTYTYRDGNNLMIQSLNADRTSLNEEBGKSVYVYNKELKRPFAV 600
QY 601 FEKPKKNAGTQNFVSIVPYDGGQKAPEISIRENKGNDFEKGLNLTITINGKQOOLVLP 659
Db 601 FEKPKKNAGTQNFVSIVPYDGGQKAPEISIRENKGNDFEKGLNLTITINGKQOOLVLP 659

RESULT 3
PCT-US95-07391A-4
; Sequence 4, Application PC/TUS9507391A
; GENERAL INFORMATION:
; APPLICANT: IBEX TECHNOLOGIES and
```

APPLICANT: ZIMMERMANN, Joseph  
 TITLE OF INVENTION: Nucleic Acid Sequences And Expression  
 TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From  
 TITLE OF INVENTION: Flavobacterium heparinum  
 NUMBER OF SEQUENCES: 26  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Hale and Dorr  
 STREET: 1455 Pennsylvania Avenue, N.W.  
 CITY: Washington, D.C.  
 COUNTRY: U.S.A.  
 ZIP: 20004  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: PCT/US95/07391A  
 FILING DATE: 09-JUNE-1995  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/258,639  
 FILING DATE: 10 JUNE 1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BAKER, Hollie L.  
 REGISTRATION NUMBER: 31,321  
 REFERENCE/DOCKET NUMBER: 104385.116PCT  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202)942-8400  
 TELEFAX: (202)942-8484  
 INFORMATION FOR SEQ ID NO: 4:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 659 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 PCT-US95-07391A-4

Query Match 100.0%; Score 3494; DB 5; Length 659;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-308;  
 Matches 659; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTTKIFKRIIVFAVIALSSGNILAOSSITRKDFDHINLEYSGLKRVKAVAGNYDDAA 60  
 Db 1 MTTKIFKRIIVFAVIALSSGNILAOSSITRKDFDHINLEYSGLKRVKAVAGNYDDAA 60  
 QY 61 KALLAYREKSKAREPDFSNAEKPADIRQPIDKVTREMAKALVHQFQPHKGYGFDYCK 120  
 Db 61 KALLAYREKSKAREPDFSNAEKPADIRQPIDKVTREMAKALVHQFQPHKGYGFDYCK 120  
 QY 121 DINWQMPVKDNEVRQQLHRVQWQAMALVYHATGDEKYAREWVYQYSDWAEKNPLGLSQ 180  
 Db 121 DINWQMPVKDNEVRQQLHRVQWQAMALVYHATGDEKYAREWVYQYSDWAEKNPLGLSQ 180  
 QY 181 DNDKFWRPLEYSRVQSLPPTFSLFVNSPAPTFPAFLMEFLNSYHQOQADYLSTHYABOQN 240  
 Db 181 DNDKFWRPLEYSRVQSLPPTFSLFVNSPAPTFPAFLMEFLNSYHQOQADYLSTHYABOQN 240  
 QY 241 HRLFEAQRNLFAGVSPFEKDSPRWRQTGISVLNTEIKKQVYADGMPQFELSPIYHVAAD 300  
 Db 241 HRLFEAQRNLFAGVSPFEKDSPRWRQTGISVLNTEIKKQVYADGMPQFELSPIYHVAAD 300  
 QY 301 IFLKAYGSAKRVNLEKEFPQSVQVQVENNIMALISISLDPDYNTPMGDSWIIDKQFRMAQ 360  
 Db 301 IFLKAYGSAKRVNLEKEFPQSVQVQVENNIMALISISLDPDYNTPMGDSWIIDKQFRMAQ 360  
 QY 361 FASWAVFPANQAIKYFATDGGKQAPNPLSKALSNAAGYTFPRSGWDMKNATVMWLKASPP 420  
 Db 361 FASWAVFPANQAIKYFATDGGKQAPNPLSKALSNAAGYTFPRSGWDMKNATVMWLKASPP 420  
 QY 421 GEFHAQPDNGTFELFTIKGRNFTPDAGVFYVSGDEAIMKLRNRYQTRIHSITLTLNQNMV 480  
 Db 421 GEFHAQPDNGTFELFTIKGRNFTPDAGVFYVSGDEAIMKLRNRYQTRIHSITLTLNQNMV 480

QY 481 ITKARQNKWETGNNLDVLTNTNPSYPNLDHORSVLFINKKYFLVIDRAIGATGNLGVHW 540  
 Db 481 ITKARQNKWETGNNLDVLTNTNPSYPNLDHORSVLFINKKYFLVIDRAIGATGNLGVHW 540  
 QY 541 QLKEDSNVFDKTKNRVVTYTRDGNLMIQSLNADRTSLNEEGKVSVYVYNKELKRPAPV 600  
 Db 541 QLKEDSNVFDKTKNRVVTYTRDGNLMIQSLNADRTSLNEEGKVSVYVYNKELKRPAPV 600  
 QY 601 FEKPKKNAGTQNFVSIYVYDQKAPESIRENKGNDFEKGLNLTTLTINGKQQLVLP 659  
 Db 601 FEKPKKNAGTQNFVSIYVYDQKAPESIRENKGNDFEKGLNLTTLTINGKQQLVLP 659

RESULT 4  
 US-09-134-000C-5302  
 ; Sequence 5302, Application US/09134000C  
 ; Patent No. 6617156  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; FILE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
 ; CURRENT APPLICATION NUMBER: US/09/134,000C  
 ; CURRENT FILING DATE: 1998-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/055,778  
 ; PRIOR FILING DATE: 1997-08-15  
 ; NUMBER OF SEQ ID NOS: 6812  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 5302  
 ; LENGTH: 649  
 ; TYPE: PRT  
 ; ORGANISM: Enterococcus faecalis  
 US-09-134-000C-5302

Query Match 4.7%; Score 164; DB 4; Length 649;  
 Best Local Similarity 20.1%; Pred. No. 5.4e-06;  
 Matches 125; Conservative 91; Mismatches 213; Indels 194; Gaps 31;

QY 120 KDINWQMPVKDNEVRQQLHRVQWQAMALVYHATGDEKYAREWVYQYSDWAEKNPLGLS 179  
 Db 70 KEYANRYPDDDPFWLMLSRQSLVLDLAQAYALTKERYLQKWSLLIDFI--NDEGSP 127  
 QY 180 QNDKFWRPLEYSRVQSLPPTFSLFVNSPAPTFPAFLMEFL-----NSYHQOQADYL 232  
 Db 128 NSTNRDWRPLDVGIRVTN-----WMKSLYIPIADPRLLGDDVLNALLIHLIDYLE 180  
 QY 233 THYAEQGNHRLFE-----AQRNLFAGVSPFEKDSPRWRQTGISVLNTEIKKQVYA 283  
 Db 181 RSYIDK--YRLSNWGVLAIGGWAIDLF---LPELVTS-KQDILWSRLAEQLDQFYS 233  
 QY 284 DGMQFELSPIYHVAADIFLKAYGSAKRVNLEKEFPQSVQVQVENNIMALISISLP---- 339  
 Db 234 DGIHWEQSPLY-----QHEVIMTFVYLLQ--ISEYLEVQLPDLIR 271  
 QY 340 -DYNTPMFGDSWITDKN-----FRMAQFASWAVFPAN 371  
 Db 272 MKLKTPIFSTHYADNQDILNPDINDSHVNFHYVDIYRKLGFIFEPSWTANRLWTGD 331  
 QY 372 -----QAIKYFATDGGKQAPNPLSKALSNAAGYTFPRSGWDMKNATVMWLKASP----- 419  
 Db 332 LYEERIWEITMK-----PKEL-----FRG-----ESSGLMAYKAEDIYFTL 366  
 QY 420 -----PGEFHAQPDNGTFELFTIKGRNFTPDAGVFYVSGDEAIMKLRNRYQTRIHSITLTL- 474  
 Db 367 FNLHGSAGHASTGCGFTLIQQLQDGLFSDSGRSYVYNKSERLQ-----KECASHNTMFTA 422  
 QY 475 DNQNVVITKARQNKWETGNNLDVLTNTNPSYPNLDHORSVLFINKKYFLVIDRAIGATG 534  
 Db 423 ENPHTLVS-----DTW-----GYDKL--PTPLFOOI-KELSVGFFAE-----456  
 QY 535 NLGVHWQLKEDSNP-VFDKTKNRVVTYTRDGNLMIQSLNADRTSLNEEGKVSVYVYN-- 591

Db 457 ---CGWLDXADQNPMIFE---RSFIYKXSINSVVIIDSPA---CQKETEITSTYNLA 504  
QY 592 -----KELKRPAFVFEKPKN---AGTQNFVSVIYPDQKAPESIRENK-----GNDF 638  
Db 505 PSINCKQEAHRFALTINKHYKTYLLFAGGQTQOSVA-----KGSEIYNQLNEHPRLSNKF 558  
QY 639 --EKGKLNLTLLTINGKQOLVNP 659  
Db 559 CYTKGREIQATVISPLEDIOITP 581

RESULT 5  
US-09-134-001C-3632  
; Sequence 3632, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 3632  
; LENGTH: 698  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-3632

Query Match 3.6%; Score 126.5; DB 4; Length 698;  
Best Local Similarity 18.6%; Pred. No. 0.015;  
Matches 139; Conservative 109; Mismatches 284; Indels 217; Gaps 31;

QY 9 IIVPAVALSSGNI-LAQSSITR--KDFHINLEYSGLEKVNKAVAAGN-----Y 56  
Db 32 VFIFAVLVLKGLYQLAQGGHYKQLIKNDENIT-----VNESVPRGILDRNGKLV 83  
QY 57 DDAKALLAYREKSKAREPDFSNAEKPAD-IRQPIDKVTREMAKALVHQFQPHKGYG 115  
Db 84 DNASKMSITYTRNKTQSKQEMLTAKKLTDLIKWDTDKITER--DK-----127  
QY 116 FDYKGINWQWP-----VKNEVFWQLHRV--KWWQA 146  
Db 128 ---KDFWIOYPSFAKKLMRKEQLMLEDSISQDQDTQLRDKIGKKQLKQTKKDLQV 183  
QY 147 MALV-----YHATGDEKYAREWVYQYSDWARKNPLGLSQDND 183  
Db 184 LAIYREMNAGSTLDPQTIKNEVDYSEKAYAVSQSLKLPVNTMTMDRKYPG---DTL 240  
QY 184 KFWRRPLEVSDRVQSIPPTSLFVNSPAFTPAFLMEFLNSYHQOADYLSHYABQGNHRL 243  
Db 241 RGIFG--DVSTSTEGIPKELT-----EQVLSKGYSRNDRVGSYLEYQVEDV 285  
QY 244 FEAQRLNLFAGVSPEFKDSPRWQGTGSLVNLTEIKQVYADGMQFELSPIYHVAADIEL 303  
Db 286 LKGTTRKQM-----KYTTDKSGRVISSVILNPGSR-----GHDLOLT-----IDIDL 326  
QY 304 KAYGSARKVN--LEKEFFPOSYQVTENIMMALISISLPDYNTPMFGDSWITDKNFMQAF 361  
Db 327 Q-----KKVESLLEKQISKLSRQAKMDNALMVQNPKNQDILATAGKQIDKQGLKQY 381  
QY 362 --ASWARRVPANQAIK-YFATDCKQKAPN-----FLSKALSNAAGFTYFRSGWKNA-----410  
Db 382 DIGNFTAQYTVGSSVKGTTLLAGYONKAINVGTMTVDEPLKFOGGLTKRSYFNKNGHVSI 441  
QY 411 -----TMVVLKASPPGCFHAQPDNGTGFELFIKGRNFTPDAGVYVYSGDE 454  
Db 442 DDQKALMHSNVSVMFKTALAGDPYTSNGSLNN-----IADAG-----481

QY 455 AIMKLRNWWYQTRIHSHTLTLDNQNMVITKARQNKWETGNMLDVLTYTNPSPNLDHORSV 514  
Db 482 --RKLKRLGNQVGLGKLTGIDLFPNETPGQIEPLTNFNGNYLDLAIGQYDTVTYPLQLSQYV 539  
QY 515 LFINKKYFLV--IDRAIGEATGNLGVHMQLKEDSNVPFDTKNRVVYTYTVDGNNLMIQS 571  
Db 540 STIANDGYRIQPHIGLSIYESTN-----KQETGPLKKIKGNVLKNVNNNSNDEKFEV 591  
QY 572 LNADRTSLNEEGKSVYVYNKELKRP-----AFVE--KPKKNAGTQNFVSVIYP---619  
Db 592 QEGFKMAFNEKQG--TGYASPRNTVWPSAGKTGTAEVFQDGEPRVNSYIGVAPVDDPKLS 650  
QY 620 ----YDQKAPESIRENKGNDFEKGLN 644  
Db 651 FSIYTYNQPVPPPWL---NGGDLGRDVIN 676

RESULT 6  
US-08-258-639A-9  
; Sequence 9, Application US/08258639A  
; Patent No. 5681733  
; GENERAL INFORMATION:  
; APPLICANT: Su, Hongsheng  
; APPLICANT: Blain, Francoise  
; APPLICANT: Bennett, Clark  
; APPLICANT: Gu, Kangfu  
; APPLICANT: Zimmermann, Joseph  
; APPLICANT: Musil, Roy  
; TITLE OF INVENTION: Nucleic Acid Sequences And Expression  
; TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From  
; TITLE OF INVENTION: Flavobacterium heparinum  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hale and Dorr  
; STREET: 1455 Pennsylvania Avenue, N.W.  
; CITY: Washington, D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/258,639A  
; FILING DATE: 10 JUNE 1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Healey, William J.  
; REGISTRATION NUMBER: 36,160  
; REFERENCE/DOCKET NUMBER: 104385.116  
; TELEPHONE: (202)942-8400  
; TELEFAX: (202)942-8484  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 23 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-258-639A-9

Query Match 3.5%; Score 124; DB 1; Length 23;  
Best Local Similarity 100.0%; Pred. No. 9.3e-05;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 414 VLKASPPGCFHAQPDNGTGFELFI 436  
Db 1 VLKASPPGCFHAQPDNGTGFELFI 23

RESULT 7  
US-08-900-951-9

```
; Sequence 9, Application US/08900951
; Patent No. 5919693
; GENERAL INFORMATION:
; APPLICANT: Su, Hongsheng
; APPLICANT: Blain, Francoise
; APPLICANT: Bennett, Clark
; APPLICANT: Gu, Kangfu
; APPLICANT: Zimmermann, Joseph
; APPLICANT: Musil, Roy
; TITLE OF INVENTION: Nucleic Acid Sequences And Expression
; TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From
; TITLE OF INVENTION: Flavobacterium heparinum
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/900,951
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/258,639
; FILING DATE: 10 JUNE 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Healey, William J.
; REGISTRATION NUMBER: 36,160
; REFERENCE/DOCKET NUMBER: 104385.116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)942 8400
; TELEFAX: (202)942 8484
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US95-07391A-9
;
; Query Match 3.5%; Score 124; DB 2; Length 23;
; Best Local Similarity 100.0%; Pred. No. 9.3e-05;
; Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 414 VLKASPPGGEFHAQPDNGTFELFI 436
; DB 1 VLKASPPGGEFHAQPDNGTFELFI 23
;
; RESULT 9
; US-08-258-639A-10
; Sequence 9, Application US/08258639A
; Patent No. 5681733
; GENERAL INFORMATION:
; APPLICANT: Su, Hongsheng
; APPLICANT: Blain, Francoise
; APPLICANT: Bennett, Clark
; APPLICANT: Gu, Kangfu
; APPLICANT: Zimmermann, Joseph
; APPLICANT: Musil, Roy
; TITLE OF INVENTION: Nucleic Acid Sequences And Expression
; TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From
; TITLE OF INVENTION: Flavobacterium heparinum
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/258,639A
; FILING DATE: 10 JUNE 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Healey, William J.
; REGISTRATION NUMBER: 36,160
; REFERENCE/DOCKET NUMBER: 104385.116
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)942 8400
; TELEFAX: (202)942 8484
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US95-07391A-9
;
; Query Match 3.5%; Score 124; DB 2; Length 23;
; Best Local Similarity 100.0%; Pred. No. 9.3e-05;
; Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 414 VLKASPPGGEFHAQPDNGTFELFI 436
; DB 1 VLKASPPGGEFHAQPDNGTFELFI 23
;
; RESULT 8
; PCT-US95-07391A-9
; Sequence 9, Application PC/TUS9507391A
; Patent No. 5919693
; GENERAL INFORMATION:
; APPLICANT: IBEX TECHNOLOGIES and
; APPLICANT: ZIMMERMANN, Joseph
; APPLICANT: Zimmermann, Joseph
; TITLE OF INVENTION: Nucleic Acid Sequences And Expression
; TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From
; TITLE OF INVENTION: Flavobacterium heparinum
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

LENGTH: 23 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-258-639A-10

Query Match 3.4%; Score 119; DB 1; Length 23;  
Best Local Similarity 91.3%; Pred. No. 0.00026;  
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 101 KALVHOFQPHKGYGFDYDKDIN 123  
Db 1 KALVHFWFPHKGYGFDYDKDIN 23

## RESULT 10

US-08-900-951-10  
Sequence 10, Application US/08900951  
Patent No. 5919693

## GENERAL INFORMATION:

APPLICANT: Su, Hongsheng  
APPLICANT: Blain, Francoise  
APPLICANT: Bennett, Clark  
APPLICANT: Gu, Kangfu  
APPLICANT: Zimmermann, Joseph

TITLE OF INVENTION: Nucleic Acid Sequences And Expression  
TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From  
TITLE OF INVENTION: Flavobacterium heparinum

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington, D.C.

COUNTRY: U.S.A.

ZIP: 20004

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC DOS/MS DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/900,951

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/258,639

FILING DATE: 10 JUNE 1994

ATTORNEY/AGENT INFORMATION:

NAME: Healey, William J.

REGISTRATION NUMBER: 36,160

REFERENCE/DOCKET NUMBER: 104385.116

TELEPHONE: (202)942 8400

TELEFAX: (202)942 8484

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 23 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-900-951-10

Query Match 3.4%; Score 119; DB 2; Length 23;  
Best Local Similarity 91.3%; Pred. No. 0.00026;  
Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 101 KALVHOFQPHKGYGFDYDKDIN 123  
Db 1 KALVHFWFPHKGYGFDYDKDIN 23

## RESULT 11

PCT-US95-07391A-10  
Sequence 10, Application PC/TUS9507391A  
GENERAL INFORMATION:

APPLICANT: IBEX TECHNOLOGIES and

APPLICANT: ZIMMERMANN, Joseph

TITLE OF INVENTION: Nucleic Acid Sequences And Expression

TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From

TITLE OF INVENTION: Flavobacterium heparinum

NUMBER OF SEQUENCES: 26

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hale and Dorr

STREET: 1455 Pennsylvania Avenue, N.W.

CITY: Washington, D.C.

COUNTRY: U.S.A.

ZIP: 20004

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/07391A

FILING DATE: 09-JUNE-1995

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/258,639

FILING DATE: 10 JUNE 1994

ATTORNEY/AGENT INFORMATION:

NAME: BAKER, Hollie L.

REGISTRATION NUMBER: 31,321

REFERENCE/DOCKET NUMBER: 104385.116PCT

TELEPHONE: (202)942-8400

TELEFAX: (202)942-8484

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 23 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

PCT-US95-07391A-10

Query Match 3.4%; Score 119; DB 5; Length 23;

Best Local Similarity 91.3%; Pred. No. 0.00026;

Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 101 KALVHOFQPHKGYGFDYDKDIN 123

Db 1 KALVHFWFPHKGYGFDYDKDIN 23

## RESULT 12

US-07-973-320-2

Sequence 2, Application US/07973320

Patent No. 5286486

GENERAL INFORMATION:

APPLICANT: Payne, Jewel M.

TITLE OF INVENTION: No. 5286486el Bacillus thuringiensis Gene

TITLE OF INVENTION: Encoding a Coleopteran-Active Toxin

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: David R. Saliwanchik

STREET: 2421 N.W. 41st Street, Suite A-1

CITY: Gainesville

STATE: FL

COUNTRY: USA

ZIP: 32606

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

;/ CURRENT APPLICATION DATA:  
;/ APPLICATION NUMBER: US/07/973,320  
;/ FILING DATE: 19921106  
;/ CLASSIFICATION: 435  
;/ PRIOR APPLICATION DATA:  
;/ APPLICATION NUMBER: US 07/788,638  
;/ FILING DATE: 6-NOV-1991  
;/ ATTORNEY/AGENT INFORMATION:  
;/ NAME: Saliwanchik, David R.  
;/ REGISTRATION NUMBER: 31,794  
;/ REFERENCE/DOCKET NUMBER: MA68.C1  
;/ TELECOMMUNICATION INFORMATION:  
;/ TELEPHONE: 904-375-8100  
;/ TELEFAX: 904-372-5800  
;/ INFORMATION FOR SEQ ID NO: 2:  
;/ SEQUENCE CHARACTERISTICS:  
;/ LENGTH: 1138 amino acids  
;/ TYPE: AMINO ACID  
;/ STRANDEDNESS: single  
;/ TOPOLOGY: linear  
;/ MOLECULE TYPE: protein  
;/ HYPOTHETICAL: YES  
;/ ANTI-SENSE: NO  
;/ ORIGINAL SOURCE:  
;/ ORGANISM: Bacillus thuringiensis  
;/ STRAIN: dakota  
;/ INDIVIDUAL ISOLATE: HDS11  
;/ IMMEDIATE SOURCE:  
;/ LIBRARY: lamdagem (TM)-11 library of J.M. Fu  
;/ CLONE: 511  
;/ US-07-973-320-2

Query Match 3.4%; Score 119; DB 1; Length 1138;  
Best Local Similarity 19.8%; Pred. No. 0.16;  
Matches 132; Conservative 85; Mismatches 206; Indels 244; Gaps 39;

Qy 48 NKAVA-----AGNYDDAAKALLAYREKSKAREPDSNAEKPADIRQPIDKVTREMAKAL 103  
Db 121 NKALAELEGLNN-----LTIYQQ--ALEDLNLPDDPATITRVIDRF--RILD-AL 167

Qy 104 VHQPHKGYGVF-----YKQINWQM-----WPVKDNEV----- 134

Db 168 FESYMSFRVAGYEILPLTVQAQANLHALLRDLSTLYGDKWGFTQNNIEENYNRQKHH 227

Qy 135 -RWQHRVKKWQA-MALVYHATGDE-----KYAREWVYQYSDWAR-----KNPLGLSQDND 183

Db 228 SEYSHCVKYNGLSRLNGSTYEQWINVNRPREMILWLDIAAVFIYDPRMYSMTS 287

Qy 184 KFWVR-----PLEVSDRVQSLPPTSLFVNSPAFTPAFLMEFLNSYHQOQADYLSHYAQ 238

Db 288 TQLTREVTDPISLSISNPDIGPFSQMENTAFRTPHLV-----DYLDLXYIT 336

Qy 239 GNHRLF--EAQRNLF-----AGVSFPEFKDSPRWOTGIVSLNTEIKQVYADGMQFELSP 292

Db 337 SKYKAFSHEIQDLEFYWCWVKHVSFKSQSNLY-TTGI-----YGKTSYSG 384

Qy 293 IYHVAADIF-----LKYGSAKRVNLEKEFFQSYVQTVENMIM 331

Db 385 AYSEFGNDIYRLAAPSVVVYPYQNYQVEFYGVKGVHYRGD--NKYDLITYDSIDQ 442

Qy 332 -----ALISISLPYNT---PMFGDSWITDKNFRMAQFASWARVFP 369

Db 443 LPDGEPIHEKYTHRLCHATAISKSTPDYDNATIFIF--SW-----THRSAEY--YNRYP 494

Qy 370 AN-----QAIKYFATGKQG--KAPNFLSKALSNAGFYTFRSGW--DKNATVMVLKASPPG 421

Db 495 NKIKKIPAVKMLDLDLSIVKPGFTGDLVKRG-----SNGYIGDIKATV---NSPLS 546

Qy 422 E-----FHAQPDNGTFELFIKGRNFTPDAGVFVYSGDEAIMKLNWVYQTRIHSTLTLD 475

Db 547 QKYRVVRVYATSVSGLFNVFI----- 567

Qy 476 NQNMWITKARQNKWET-GNNLDVITY-----TNPSYPN-----LDH--QRSVL 515

Db 568 NDEIALQKNEFQSVETIGCGKD-LTYGSFGYEVSTTIQFPNEHPKTLHLNHLNNSPF 626

Qy 516 FINKKYELVIDRAIGEATGNLGVHWQLKEDSNPVDFTKRVYTYTTRDGNLM---IQSL 572

Db 627 YVDSIEFIPVD-----VNYDEKEK-----LEKAQKAVNTLFTTEGRNALQKYVTDY 671

Qy 573 NADRTSL 579

Db 672 KVDQVSI 678

RESULT 13

US-07-973-320-4

;/ Sequence 4, Application US/07973320

;/ Patent No. 5286486

;/ GENERAL INFORMATION:

;/ APPLICANT: Payne, Jewel M.

;/ APPLICANT: Fu, Jenny M.

;/ TITLE OF INVENTION: No. 5286486el Bacillus thuringiensis Gene

;/ TITLE OF INVENTION: Encoding a Coleopteran-Active Toxin

;/ NUMBER OF SEQUENCES: 4

;/ CORRESPONDENCE ADDRESS:

;/ ADDRESSEE: David R. Saliwanchik

;/ STREET: 2421 N.W. 41st Street, Suite A-1

;/ CITY: Gainesville

;/ STATE: FL

;/ COUNTRY: USA

;/ ZIP: 32606

;/ COMPUTER READABLE FORM:

;/ MEDIUM TYPE: Floppy disk

;/ COMPUTER: IBM PC compatible

;/ OPERATING SYSTEM: PC-DOS/MS-DOS

;/ SOFTWARE: Patent In Release #1.0, Version #1.25

;/ CURRENT APPLICATION DATA:

;/ APPLICATION NUMBER: US/07/973,320

;/ FILING DATE: 19921106

;/ CLASSIFICATION: 435

;/ PRIOR APPLICATION DATA:

;/ APPLICATION NUMBER: US 07/788,638

;/ FILING DATE: 6-NOV-1991

;/ ATTORNEY/AGENT INFORMATION:

;/ NAME: Saliwanchik, David R.

;/ REGISTRATION NUMBER: 31,794

;/ REFERENCE/DOCKET NUMBER: MA68.C1

;/ TELECOMMUNICATION INFORMATION:

;/ TELEPHONE: 904-375-8100

;/ TELEFAX: 904-372-5800

;/ INFORMATION FOR SEQ ID NO: 4:

;/ SEQUENCE CHARACTERISTICS:

;/ LENGTH: 1138 amino acids

;/ TYPE: AMINO ACID

;/ STRANDEDNESS: single

;/ TOPOLOGY: linear

;/ MOLECULE TYPE: protein

;/ HYPOTHETICAL: YES

;/ ANTI-SENSE: NO

;/ ORIGINAL SOURCE:

;/ ORGANISM: Bacillus thuringiensis

;/ STRAIN: kumamotoensis

;/ INDIVIDUAL ISOLATE: HD867

;/ IMMEDIATE SOURCE:

;/ LIBRARY: lamdagem (TM)-11 library of J.M. Fu

;/ CLONE: 867

;/ US-07-973-320-4

Query Match 3.4%; Score 119; DB 1; Length 1138;  
Best Local Similarity 19.8%; Pred. No. 0.16;  
Matches 132; Conservative 85; Mismatches 206; Indels 244; Gaps 39;

Qy 48 NKAVA-----AGNYDDAAKALLAYREKSKAREPDSNAEKPADIRQPIDKVTREMAKAL 103

Db 121 NKALAELEGLNN-----LTIYQQ--ALEDLNLPDDPATITRVIDRF--RILD-AL 167

QY 104 VHQFQPHKGYGYP-----YKXDNWQM-----WPKDNEV----- 134  
Db 168 PESIMPSFRVAGYEIPLLTYAQAHLALLDSTLYGDKWGTQNNIENYNRQKKHI 227  
QY 135 -RWQLHRVKKWQA-MALVYHATGE-----KYAREWVYQYSDWAR-----KNPLGLSDND 183  
Db 228 SEYSNCHVKWYNSGLSLNGSTYEQWYNRFRREMILMVLDAAPPIYDPRMYSNETS 287  
QY 184 KFWR-----PLEVDRVQSLPTFTSLFVNSPAPFTAPLMEFLNSYHQOADIYLSHYAEQ 238  
Db 288 TQLTREYVTPDISLISNPDIGSFQSMNTAFTPLV-----DYLDLYIYT 336  
QY 239 GNRHLP--EAQRNLF-----AGVSPPEPKSPRWQTCISVLNTEIKQVADGMQFELSP 292  
Db 337 SKYKASHEIQPDLFYWCVHKVFKKSEQNLY--TTGI-----YKTSYGISSG 384  
QY 293 IYHVAARDIF-----LKAYGSAKRVNLEKEFPQSYVQTVENMIM 331  
Db 385 AYSFRGNDIVRTLAAPSVVVYPYQNYGVQEVFYGKGVHYRGD--NKYDLTYDSIDQ 442  
QY 332 -----ALISISLPDNT-----PMFGDSWITDKNPRMAQFASWARVFP 369  
Db 443 LPPDGEPIHEKYTHRLCHATAISKSTPDYDNATIPF--SW-----THRSAEY--YNRiYP 494  
QY 370 AN---QAIKYEATDGKQG--KAPNLSKALSNAAGFYTFRSGW--DKNATVMVLKASPPG 421  
Db 495 NKIKKIPAVKQYKLDLSTVKGPGFTGGDLVVRG-----SNGYIGDITATV-----NSPLS 546  
QY 422 E-----FHAQPDNGTFELFKGRNFTPDAGVFVYSGDEAIMKLRNMYRQTRIHSITLTD 475  
Db 547 QKZRVVRVATSVSGLFNVI-----VNYDEKEK-----LEKAQAVNTLFTGGRNALQKYTDY 671  
QY 476 NQNMVITKARQWET--GNLDVITY-----TNPSYPN-----LDH--QRSVL 515  
Db 568 NDBIALQKPNQSTVETIGEGK-LTYGSGFYIEYSTIOPNPHPKITLHLNLSNNSPF 626  
QY 516 FINKKYFLVIDRAIGATGNLGVHWOLKEDSNFVFKTKURVYTVYRDGNLNM--IQSL 572  
Db 627 YVDSIEFIPVD-----VNYDEKEK-----LEKAQAVNTLFTGGRNALQKYTDY 671  
QY 573 NADRTSL 579  
Db 672 KVDQVSI 678

## RESULT 14

US-08-258-639A-2  
; Sequence 2, Application US/08258639A  
; Patent No. 5681733  
; GENERAL INFORMATION:  
; APPLICANT: Su, Hongsheng  
; APPLICANT: Blain, Francoise  
; APPLICANT: Bennett, Clark  
; APPLICANT: Gu, Kangfu  
; APPLICANT: Zimmermann, Joseph  
; APPLICANT: Musil, Roy  
; TITLE OF INVENTION: Nucleic Acid Sequences And Expression  
; TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From  
; TITLE OF INVENTION: Flavobacterium heparinum  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hale and Dorr  
; STREET: 1455 Pennsylvania Avenue, N.W.  
; CITY: Washington, D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/258,639A  
; FILING DATE: 10 JUNE 1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Healey, William J.  
; REGISTRATION NUMBER: 36,160  
; REFERENCE/DOCKET NUMBER: 104385.116  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)942-8400  
; TELEFAX: (202)942-8484  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 772 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-258-639A-2

Query Match 3.4%; Score 118; DB 1; Length 772;

Best Local Similarity 20.5%; Pred. No. 0.11;  
Matches 128; Conservative 79; Mismatches 229; Indels 188; Gaps 30;

QY 145 QAMALVYHATGDEKYAREWVYQYSD---WARKNPLG--LSQDNDFVWRPLEVSDRV--QS 198  
Db 109 ELMALNVLMTKDPKVGREAITSIIDLETATFPAGDISRGIGLFWVTGAIVDWCYDQL 168  
QY 199 LPPTFSLFVNS-----PAPTPAFMEFLNSYHQOADIYLSHYAEQGNHRLFEA 246  
Db 169 KPEKTRFRVKAFLAKMLECGYPPVKDKSIVGHASEWMIMRDLISVGLIAYDE---FPE 225  
QY 247 QRNLFAGVSPPEPKDSPRW-----RQTGISVLNTEIKQVYA-----DGM 286  
Db 226 MYNLAAGRFKEHLVARNFYPSHNYHQGMSYLVNFTNDLFALWILDRMGAGNVNPGQ 285  
QY 287 QFELSPYIYHAAIDIFLKAYGSAKRVNLEKEFPQSYVQTVENMIMAILISLDPYNTPMF 346  
Db 286 QFTLYDAIKRRPDGQILAGGD---VDYSRKPKYIT---MPALLAGSY--YKDEYL 334  
QY 347 GUSWITDKNPRMAQFASWARVFPANQAIKYFATDGKQG--KAPNLSKAL---SNAGPYTF 402  
Db 335 NYEFLKDPN-----VEPHCKLFELMRDTQLGSRKPDLDPLSRYSYSGSPFGWMIA 383  
QY 403 RSGWCKNATVMVLKASPPGCF-----HAQPDNGTFELFKGRNFTPDAGVVF--YSGDEAIM 457  
Db 384 RTGWGPESVIAEMKN---EYSLNHHQHDAGAFQIYYKG--PLAIDAGSYTSGSSGVNSP 439  
QY 458 KLRNMYRQTRIHSITLTL-----DNQNMVITKARQ----- 486  
Db 440 HKNKPFKRTIAHNSLLIYDPKETFFSSSGYGGSDHTDFAANDGGQRLPGKGIAPRLKEM 499  
QY 487 --NKWETGNL-----DVLTYNPSYPNL-----DHORSVLFINKK----- 520  
Db 500 LAGDFRTGKILAQGGFDPNQT---PDYTLKGDITAAYSKAKVEKRSFLMLKDAKVP 556  
QY 521 -YFLVIDRAIGATGNLGVHWOLKEDSNFVFK-----TKNRVTVTVYRDGN-- 566  
Db 557 AAIVFDKVA-----SNPDFKFKLLHSIEQPEIKGNQITIKRTKNGDS 601  
QY 567 -----LMQSLNADRTSLNEEBGKVSYY-----NKLKRPAPVFE--KPKK 606  
Db 602 GMLVNTALLPDAANSNITSIG--GKGDFFWVFGTNTDPKGTDEALERGERVETPKK 660  
QY 607 NAGTQNFVSIYVYDQ--OKAPEI 628  
Db 661 AAAEDYLVNVIQIADNTQOKLHEV 684

## RESULT 15

US-08-900-951-2  
; Sequence 2, Application US/08900951  
; Patent No. 5919693  
; GENERAL INFORMATION:  
; APPLICANT: Su, Hongsheng



APPLICANT: Blain, Francoise  
APPLICANT: Bennett, Clark  
APPLICANT: Gu, Kangfu  
APPLICANT: Zimmermann, Joseph  
APPLICANT: Musil, Roy  
TITLE OF INVENTION: Nucleic Acid Sequences And Expression  
TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From  
TITLE OF INVENTION: Flavobacterium heparinum  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hale and Dorr  
STREET: 1455 Pennsylvania Avenue, N.W.  
CITY: Washington, D.C.  
COUNTRY: U.S.A.  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC DOS/MS DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/900,951  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/258,639  
FILING DATE: 10 JUNE 1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Healey, William J.  
REGISTRATION NUMBER: 36,160  
REFERENCE/DOCKET NUMBER: 104385.116  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)942 8400  
TELEFAX: (202)942 8484  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 772 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-900-951-2

Query Match 3.4%; Score 118; DB 2; Length 772;  
Best Local Similarity 20.5%; Pred. No. 0.11;  
Matches 128; Conservative 79; Mismatches 229; Indels 188; Gaps 30;  
QY 145 QAMALVYHATGKXAREWVYQYSD---WARKNPLG-LSQDNDKFWRPLEVSDRV--QS 198  
DB 109 ELMALNYLMTKPKVGREAITSIIDTLETATFKPAGDISRGIGLEFWVTGAIYVDWCYDQL 168  
QY 199 LPPTFSLFVNS-----PATTAPFLMEFLNSHQADYLSHYABQGNHRLPEA 246  
DB 169 KPEEKTRFVKAFVRLAKMLECGYPPVKOKSIVGHASEWMIMRDLLSVGAIYDE---PPE 225  
QY 247 QNLFAGVSFPEFKDSPW-----RQTGISVLNTEIKKQVVA-----DGM 286  
DB 226 MYNLAAGRFFKEHLVARNWFPYSHNVHQGWSVLNVRFTNDLFALWILDRMGAGNVFNPOQ 285  
QY 287 QPELSPYHVAADIFLKAYSKARVLEKEFPQSVQTVENMIMALISISLPDYNTPMF 346  
DB 286 QFILDYAIYKRRPDGQILLAGD---VDYSRKKPKYIT-----MPALLAGSY--YKDEYL 334  
QY 347 GDSWITDKNFRMAQFASWAFVPAQAIKYFATDGKQ--KAPNLSKAL---SNAGFYTF 402  
DB 335 NYEFLKDPN-----VEPHCKLFEFLWRDTQLGSRKPDLLPLSRYSGSPFGWMIA 383  
QY 403 RSGWKNATVMVLKASPPGEF---HAQPDNGTFFELFKGRNFTPDAGVFV--YSGDEAIM 457  
DB 384 RTGWGPESVIAEMKVN---EYSLFNHQHQDAGAFQIYYKG--PLAIDAGSYTSSGGYNP 439  
QY 458 KLRNMYRQTRIHSITLTL-----DNQNMVITKARQ----- 486  
DB 440 HKNKFFKRTIAHNSLLIYDPKETSSSGYGGSDHTDFAANDGQRLPGKGIAPRLDKEM 499

QY 487 --NKWETGNL-----DVLTYTNPSPYNL-----DHORSVLPIKK----- 520  
DB 500 LAGDFRTGKILAQGFDPNQ---PDYTLKGDITAAYSAKVKEVRSFLNLLKDAKVP 556  
QY 521 -YFLVIDRAIGEATGNLGVHQLKEDSNPVDK-----TKNRVYVYTYRDGNN-- 566  
DB 557 AAMIVFDKVA-----SNPDKKFWLLHSIEQPEIKGNQITIKRTKNGDS 601  
QY 567 -----LMIQSLNADRTSLNEEEKVSYVY-----NKKELKRPAPVFE-KPKK 606  
DB 602 GMLVNTALLPDAANSNITSIG-GKGKDFWVGTNTYNDPKPTDEALERGEWRVEITPK 660  
QY 607 NAGTONFVSIVPYDG--OKAPEI 628  
DB 661 AAEDYLNVIQIADNTQOKLHEV 684

Search completed: July 27, 2004, 17:32:01  
Job time : 22 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 31, 2004, 15:56:51 ; Search time 993 Seconds  
(without alignments)  
3253.947 Million cell updates/sec

Title: US-09-802-285A-2

Perfect score: 3494

Sequence: 1 MTTKIFKRIIVFAVIALSSG.....KGLNLTITNGKQQLVLP 659

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3222919 seqs, 2451570024 residues

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool.p/US09802285/runat\_27072004\_155105\_9097/app.query.fasta\_1.839  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09802285 @cgn 1 1 912 @runat 27072004 155105\_9097  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/2/pubpna/PTC\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:  
6: /cgn2\_6/ptodata/2/pubpna/PTUS\_PUBCOMB.seq:  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:  
18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

ALIGNMENTS

RESULT 1  
US-10-282-122A-9400  
; Sequence 9400, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21

1	243	7.0	2688	13	US-10-282-122A-9400	Sequence 9400, Ap
2	176	5.0	3763	13	US-10-158-844-186	Sequence 186, App
c	166.5	4.8	11427	9	US-09-070-927A-165	Sequence 165, App
4	139	4.0	2295	16	US-10-398-221-2467	Sequence 2467, Ap
c	139	4.0	3011208	16	US-10-398-221-2058	Sequence 2058, Ap
6	132.5	3.8	2088	13	US-10-282-122A-34562	Sequence 34562, A
7	130.5	3.7	2043	13	US-10-282-122A-35377	Sequence 35377, A
8	123	3.5	1599	16	US-10-369-493-45761	Sequence 45761, A
9	122.5	3.5	3753	10	US-09-769-736-9	Sequence 9, Appli
10	121	3.5	2541	16	US-10-320-797-2302	Sequence 2302, Ap
11	120	3.4	2034	16	US-10-369-493-46899	Sequence 46899, A
12	118.5	3.4	1737	16	US-10-369-493-26008	Sequence 26002, A
13	118.5	3.4	1737	16	US-10-369-493-26008	Sequence 26008, A
14	118	3.4	7787	16	US-10-398-221-3883	Sequence 3883, Ap
15	117.5	3.4	4455	13	US-10-335-977-4204	Sequence 4204, Ap
16	117.5	3.4	8709	13	US-10-282-122A-22858	Sequence 22858, A
17	116	3.3	2757	13	US-10-282-122A-34553	Sequence 34553, A
c	116	3.3	4702	8	US-08-781-986A-268	Sequence 268, App
c	116	3.3	4702	13	US-10-329-824-268	Sequence 268, App
19	116	3.3	14121	10	US-09-802-840-31	Sequence 31, Appli
20	116	3.3	14121	10	US-09-920-033-3	Sequence 31, Appli
21	116	3.3	14121	10	US-10-403-902A-31	Sequence 31, Appli
22	116	3.3	14121	13	US-10-147-196-3	Sequence 3, Appli
23	116	3.3	14121	16	US-10-388-263-545	Sequence 545, App
24	116	3.3	2961	13	US-10-282-122A-12261	Sequence 12261, A
25	115.5	3.3	3945	16	US-10-369-493-25219	Sequence 25219, A
26	115.5	3.3	3945	16	US-10-369-493-25219	Sequence 369, App
27	115.5	3.3	3996	9	US-09-801-368-369	Sequence 10, Appl
28	115.5	3.3	1163020	16	US-10-398-221-10	Sequence 2058, Ap
29	115.5	3.3	3011208	16	US-10-398-221-2058	Sequence 2654, Ap
30	115	3.3	3222	9	US-09-974-300-2654	Sequence 74, Appli
31	115	3.3	16535	13	US-10-158-844-74	Sequence 1, Appli
c	115	3.3	1830121	15	US-10-329-960-1	Sequence 1, Appli
c	115	3.3	1830121	16	US-10-329-960-1	Sequence 387, App
33	114.5	3.3	3903	10	US-09-882-227-387	Sequence 1, Appli
34	114.5	3.3	1830121	15	US-10-329-960-1	Sequence 1, Appli
35	114.5	3.3	1830121	16	US-10-329-960-1	Sequence 37158, A
36	114.5	3.3	1830121	16	US-10-329-960-1	Sequence 10381, A
37	114	3.3	2613	16	US-10-369-493-37158	Sequence 7, Appli
38	114	3.3	6702	13	US-10-282-122A-10381	Sequence 5, Appli
39	113.5	3.2	4765	14	US-09-841-553-7	Sequence 321, App
40	113.5	3.2	4765	14	US-10-090-624-5	Sequence 41762, A
41	112.5	3.2	6720	9	US-09-070-927A-321	Sequence 34629, A
42	111.5	3.2	2889	13	US-10-282-122A-41762	Sequence 19, Appl
43	111	3.2	2169	16	US-10-369-493-34629	Sequence 18, Appl
44	111	3.2	2253	9	US-09-993-292A-18	
45	111	3.2	6908	9	US-09-993-292A-18	

; PRIOR APPLICATION NUMBER: 60/206,848  
 ; PRIOR FILING DATE: 2000-05-23  
 ; PRIOR APPLICATION NUMBER: 60/207,727  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 60/230,335  
 ; PRIOR FILING DATE: 2000-09-06  
 ; PRIOR APPLICATION NUMBER: 60/230,347  
 ; PRIOR FILING DATE: 2000-09-09  
 ; PRIOR APPLICATION NUMBER: 60/242,578  
 ; PRIOR FILING DATE: 2000-10-23  
 ; PRIOR APPLICATION NUMBER: 60/253,625  
 ; PRIOR FILING DATE: 2000-11-27  
 ; PRIOR APPLICATION NUMBER: 60/257,931  
 ; PRIOR FILING DATE: 2000-12-22  
 ; PRIOR APPLICATION NUMBER: 60/267,636  
 ; PRIOR FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: 60/269,308  
 ; PRIOR FILING DATE: 2001-02-16  
 ; Remaining Prior Application data removed - See File Wrapper or PALM.  
 ; NUMBER OF SEQ ID NOS: 78614  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 9400  
 ; LENGTH: 2688  
 ; TYPE: DNA  
 ; ORGANISM: Bacillus anthracis  
 ; US-10-282-122A-9400

## Alignment Scores:

Pred. No.: 4,83e-17 Length: 2688  
 Score: 243.00 Matches: 132  
 Percent Similarity: 37.82% Conservative: 93  
 Best Local Similarity: 22.18% Mismatches: 250  
 Query Match: 6.95% Indels: 120  
 DB: 13 Gaps: 29

US-09-802-285A-2 (1-659) x US-10-282-122A-9400 (1-2688)

QY	113	TyrGlyTyrPheAspTyrGlyLysAspIleAsnTyrGlnMetTrpProValLysAspAsn	132
Db	139	TTTCGGAAGTAGTTTCATATCTTAATTAATTAATCTGG-----AAT	177
QY	133	GluValArgTrpGln-----LeuHisArgValLysTrpTrpGlnAlaMet	147
Db	178	GAACAAAGAGACGGGGATATTTAAGATTAAATACACGGTTCATACGTTTTCAGCATGCTTA	237
QY	148	AlaLeuValTyrHisAlaThrGlyAspGluLysTyrAlaArgGluTrpValTyrGlnTyr	167
Db	238	ATAGCCCGTATAATGATCTGGTGATGATGAAGATATATAAAATATATAAAATCTATAGAGTTAAAT	297
QY	168	SerAspTrpAlaArgLysAsnProLeuGlyLeuSerGlnAspAsnAspLysPheValTyr	187
Db	298	AAAGATTGGATTAAATATCATTCATTTGAACITTCACCAACATTCATGCTTT-----351	
QY	188	ArgProLeuGluValSerAspArgValGlnSerLeuProThrPheSerLeuPheVal	207
Db	352	-----CATGATGAACTACAGCGCTA-----CGTCTACAGTATTGGTTA	390
QY	208	AsnSerProAlaPheThrProAlaPheLeuMetGlu-----PheLeuAsn	222
Db	391	AGATTCTACATTTTACAGCTCAAGTGTTCATCAGGGAAGAATATATTTGTAGAAAAA	450
QY	223	SerTyrHisGlnAlaAspTyrLeuSerThrHisTyr-----AlaGluGlnGlyAsn	240
Db	451	AGTATGGAAGATACTCGCAAACTTTTATCAGAGGACTTTTTCATGCTACTAATAACAAT	510
QY	241	HisArgLeuPheGluAlaGlnArgAsnLeuPheAlaGlyValSerPheProGluPheLys	260
Db	511	CATGGTATGTTTCAA---GATAGAGCGCTACTTACTTATGTCATCGTAT-----TTTAAA	561
QY	261	---AspSerPro-----ArgTrpArgGlnThrGlyLeuSerValLeuAsnThrGlu	276
Db	562	GGAGAAATCCCTCCCTTGAAAAATATATAAAGTTGGCTGTACACGTTTAAAGATTAT	621
QY	277	IleLysLysGlnValTyrAlaAspGlyMetGlnPheGluLeuSerProIleTyrHisVal	296
Db	622	TTTGAAAAGGTCTTTTACTGAAGAAGGGGTTTCATAAGGAGCATTCACCATCTTATCATTTA	681
QY	297	AlaAlaIleAspIlePheLeuLysAlaTyrGlySerAlaLysArgValAsnLeuGluLys	316
Db	682	-----TTAGTAGCTTCTAATATCAAGAAGTTAGCAAACTGGATGAAA	723
QY	317	GluPheProGlnSerTyrValGlnThrValGluAsnMetIleMetAla-----	332
Db	724	GAATTCGATAAGCAA---GTAAGTTTCATTTTAATAAGATTATATAAAAAACAGAAGAG	780
QY	333	---LeuIleSerIleSerLeuProAspTyrAsnThrProMetPheGlyAspSerTrp---	350
Db	781	TATGCCATACATATCATTCCTCCAGATGGTCTTTACCACCAATATGTGATACGGAAGCA	840
QY	351	---IleThrAspLysAsnPheArgMetAlaGlnPheAlaSerTrpAlaArgValPhePro	369
Db	841	AATTTACTGGGAATAATTTATAAG-----GATTATATGAG	876
QY	370	AlaAsnGlnAlaIleLysTyrPheAlaThrAspGlyLysGlnGlyLysAlaProAsnPhe	389
Db	877	AGTGATCAGTATCTT---TATGTAGTTACTAAGGCAAAAAGGAAAGCCCTACAGAA	933
QY	390	LeuSerLysAlaLeuSerAsnAlaGlyPheTyrThrPheArgSerGlyTrpAspLys---	408
Db	934	GATGATAAGGTATTTCTCAATCAGGTATGCAATCTTTAGGAATGATTCGAGTAAGAA	993
QY	409	---AsnAlaThrValMetValLeuLysAlaSerProGlyGluPheHisAlaGlnPro	427
Db	994	GAGAAGCTACTTATGTTTATTACAGCTGCTTATCATCATAGATTACCATATAGCATAGT	1053
QY	428	AspAsnGlyThrPheGluLeuPheIleLysGlyArgAsnPheThrProAspAlaGlyVal	447
Db	1054	GAT-----GATTTAAATTTA	1068
QY	448	PheValTyrSerGlyAspGluAlaIleMetLysLeu---ArgAsnTrpTyrArg-----	464
Db	1069	TATATTTCATCAATCGTGAATTTACAGAAGCGGGTCTTAATAGATATAATTTATAA	1128
QY	465	-----GlnThrArgIleHisSerThrLeuThrLeuAspAsn	476
Db	1129	GATCCTTTTACTGAATATCGTATTTCATCTATTGCTCATATAACATCATGTTGTTCGGA	1188
QY	477	GlnAsnMetValIleThrLysAlaArgGlnAsnLys-----TrpGluThr	491
Db	1189	AAAGGGTTACCTAGAACATCATCGTCAGTAGTAAAGAAAGTCTATTATCAGATTATGAATC	1248
QY	492	GlyAsnAsnLeuAspValLeuThrTyrThrAsnProSerTyrProAsnLeuAspHisGln	511
Db	1249	AACAAGATAAAGTTGAACGCGGGAATTAATTTACGTTTATACAGGTGTGACATAGT	1308
QY	512	ArgSerValLeuPheIleAsnLysLysTyrPheLeuValIle---AspArgAlaIleGly	530
Db	1309	CGTACTGTAAAGTTACATGAAGACGAAGAAAAATAGTAGTAGGAGATTGGTGAATCT	1368
QY	531	GluAlaThrGlyAsnLeuGlyValHisTrpGlnLeuLysGluAspSerAsnProValPhe	550
Db	1369	GATAAACGTCATGAATATTAAGTTTATATGGCATGTTGTCATCAGAT-----ATTACA	1419
QY	551	AspLysThrLysAsnArgValTyrThrTyrArgAspGlyAsnLeuMetIleGln	570
Db	1420	GTACATGTAAGAGATCGTATTCTTGAGTTATTTCTGTACACCACTAAAGATATGGAATG	1479
QY	571	SerLeu-----AsnAlaAspArgThrSerLeuAsnGluGlu---GluGlyLys	585
Db	1480	GAAGTAACCTACAGTTTACTGGTGTATATACGAGCATTAATATGAAACAAACAAAGCCTCAG	1539
QY	586	ValSerTyrValTyrAsnLysGluLeuLysArgProAlaPheValPheGluLysProLys	605
Db	1540	GTGAGT-----GGCTGGGTATTTCCAAAAATGGGG	1569
QY	606	LysAsnAlaGlyThrGlnAsnPheValSerIleValTyrProTyrAspGlyGlnLysAla	625

Db 1570 GAAAAACGAGTGCT-----ACAAATGAGGTGGACATTAGGTGTTCTAATGTA 1620  
Qy 626 ProGluIleSerIleArgGluAsnLysGlyAsnAspPheGluLysGlyLys----- 642  
Db 1621 GAG-----TGTAACAGAGTTTCGATTGAAGGATTTTAAATAGGTAGACAGATTG 1674  
Qy 643 -----LeuAsnLeuThrLeuThrIleAsnGlyLysGlnGlnLeu 655  
Db 1675 ATTCCTCTAATTTTAGAGAAAACATTCAGAGAGTACACGTAATTTA 1719  
RESULT 2  
US-10-158-844-186/c  
; Sequence 186, Application US/10159844  
; Publication No. US20040029118A1  
; GENERAL INFORMATION:  
; APPLICANT: Kunsch et al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-R  
; COMPUTER: Dell Latitude Pentium 3  
; OPERATING SYSTEM: Windows 98  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/158,844  
; FILING DATE: 03-Jun-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/961,527  
; FILING DATE: 1997-10-30  
; APPLICATION NUMBER: US 60/029,960  
; FILING DATE: 1996-10-31  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hyman, Mark J.  
; REGISTRATION NUMBER: 46,789  
; REFERENCE/DOCKET NUMBER: PB340PID1  
; INFORMATION FOR SEQ ID NO: 186:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3763 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 186:  
US-10-158-844-186  
Alignment Scores:  
Pred. No.: 4.8e-09 Length: 3763  
Score: 176.00 Matches: 116  
Percent Similarity: 34.63% Conservative: 80  
Best Local Similarity: 20.49% Mismatches: 231  
Query Match: 5.04% Indels: 140  
DB: 13 Gaps: 20  
US-09-802-285A-2 (1-659) x US-10-158-844-186 (1-3763)  
Qy 71 SerLysAlaArgGluProAspPheSerAsnAlaGluLysProAlaAspIleArgGlnPro 90  
Db 2173 TCAAAAGACTATGAGAAAGTCAAGAGTTCCCTTGAACGCTTGATGACAAATCGTTTATG 2114  
Qy 91 IleAspLysValThrArgGluMetAlaAspLysAlaLeuValHisGlnPheGlnProHis 110  
Db 2113 TTGTATAGT---CCTTGGGATATGGAGCCTTGTTCAAAAATCCATCAAAATCCAGCG--- 2060  
Qy 111 LysGlyTyrGlyTyrPheAspTyrGlyLysAspIleAsnTyrGlnMetTrpProValLys 130

Db 2059 -----ATGGTATGGGATCAAGTATTTGAAGAT 2033  
Qy 131 AspAsnGluValArgTrpGlnLeuHisArgValLysTrpTrpGlnAlaMetAlaLeuVal 150  
Db 2032 GATCCAGAATGGTCTTATATGCTCAATCGACAAGAATATCTCTTGCAGTATTATGATAGGG 1973  
Qy 151 TyrHisAlaThrGlyAspGluLysTyrAlaArg-----GluTrp 163  
Db 1972 TATCTGTAGAGGAGAGATAAGGACTATATCAAAAGTCAAGTCTCTTTCTATTTGATTGG 1913  
Qy 164 ValTyrGlnTyrSerAspTrpAlaArgLysAsnProLeuGlyLeuSerGlnAspAsnAsp 183  
Db 1912 ATTGAGCAGGTGAGAGAATTTTCTCCTCAATCCTTGATGACTAGAACCTTGATAGCGGT 1853  
Qy 184 -----LysPheValTrpArgProLeuGluValSerAspArgValGlnSerLeuProPro 201  
Db 1852 ATTCGTTCTTTACTTGTGTTGAAACTA----- 1826  
Qy 202 ThrPheSerLeuPheValAsnSerProAlaPheThrProAlaPheLeuMetGluPheLeu 221  
Db 1825 ---CTCTTGTCTCTCTCTGAAATTTGACTGTGTAGAGGAGAAAGAACTAGAGAAATTTG 1769  
Qy 222 AsnSerTyrHisGlnGlnAlaAspTyrLeuSerThrHisTyrAlaGluGln----- 238  
Db 1768 GTCACTAGAAAAGCAGATTGACTTTATGAAAAGCTACTATCGCCCAAGTACACCTT 1709  
Qy 239 GlyAsnHisArgLeuPheGluAlaGlnArgAsnLeuPheAlaGlyValSerProGlu 258  
Db 1708 AGTAACTGGGGATTTTACAAACAATTCGATGCTTGTCTATCTATCAITCTTTTCAGAT 1649  
Qy 259 PheLysAspSerProArgTrpArgGlnThrGlyIleSerValLeuAsnThrGluLys 278  
Db 1648 AAGATGGATCTAGAAGAAGCTTACCATTTGCTTCAGAGAGTTGAAACAGCAAAATGAG 1589  
Qy 279 LysGlnValTyrAlaAspGlyMetGlnPheGluLeuSerProIleTyrHisVal----- 296  
Db 1588 ACACAGATTTTAGGAGATGGAAGCCAGTTTGAACAGTCGATTCTCTATCATGTAGAGTT 1529  
Qy 297 -----AlaAlaIleAspIlePheLeuLysAlaTyrGlySerAlaLysArgValAsnLeu 314  
Db 1528 TATAAAGCCTTGCTGCAATTTGTGCTC-----TTG 1499  
Qy 315 GluLysGluPheProGlnSerTyrValGlnThrValGluAsnMetIleMetAlaLeu 334  
Db 1498 CTTCCAGACTTGCAGAGATAGTTTACCAAGATGCTGCGAAAGATGCGCACTATTATCAA 1439  
Qy 335 SerIleSerLeuProAspTyrAsnThrProMetPheGlyAspSerTrpIleThrAspLys 354  
Db 1438 ATGATGACAGCGCTTAGATGGACGACTTTGGCTTTTGGTGTAGCGATTCTACAGAAACG 1379  
Qy 355 AsnPheArgMetAlaGlnPheAla----- 362  
Db 1378 ACAGAAATTTTGAGCCTGTCTGCTGTTTGAACACGAGAGACCTTCTTAAACGGTCTG 1319  
Qy 363 -----SerTrpAlaArgValPheProAlaAsnGlnAlaIleLysTyrPhe 377  
Db 1318 GATGTTAAAGTTGATTTGCTTAGCCTCTTGTTCTCGGGCGGAGAAAGTCAAGCGACTG 1259  
Qy 378 AlaThrAspGlyLysGlnGlyLysAlaProAsnPheLeuSerLysAlaLeuSerAsnAla 397  
Db 1258 CAGGAATTTGAAAAGAGAGCTTGGCAGCCTAAG-----TCCATGATCTTTTGAAGACTCT 1205  
Qy 398 GlyPheTyrThrPheArgSerGlyTrpAspLysAsnAlaThrValMetValLeuLysAla 417  
Db 1204 GGACATGTCTGCATTAAAGGAT-----GAACATCGTTATCTATTTTCAAAAT 1157  
Qy 418 SerProGlyGluPheHisAlaGlnProAspAsnGlyThrPheGluLeuPheLys 437  
Db 1156 GGTCCGCTAGGAGTCCCAATAGTCATAGTACAGAGATAGTTTTTGTTCAGATATCAA 1097  
Qy 438 GlyArgAsnPheThrProAspAlaGlyValPheValTyrSerGlyAspGluAlaLeuMet 457  
Db 1096 GSCCAACCTATTTTCATAGATCTGGCGTTTATCTTATCG--GGAGATATATGAACGTTA 1038

```
QY 458 LysLeuArgAsnTrpTyrArgGlnThrArgIleHisSerThrLeuThrLeuAspAsnGln 477
Db 1037 TCCTCTAAAGAGCTGGAGTCAATCGACTGCATTTG----- 999
QY 478 AsnMetValIleThrLysAlaArgGlnAsnLysTrpGlu--ThrGlyAsnAsnLeuAsp 496
Db 998 -----AGATGGGAAGAGCTCCGGAAGATCACGGG 969
QY 497 ValLeu-ThrTyr----- 500
Db 968 ATCTGTGGAATATGAATACTATCTCCTCCTGTTTGTCCACCATAAAGAGGAGGG 909
QY 501 -----ThrAsnProSerTyrProAsnLeuAspHi 510
Db 908 AGTGCAATATTGAGGGGGCTTATGTGTGAGCAAGAACCTGATTGCTTATCTTT---CA 852
QY 510 sGlnArgSerValLeuPheIleAsnLysLysTyrPheLeuValIleAsp-----ArgAl 528
Db 851 CAGGAGAAATCTCATGTTGGTAGAGGATGCTGGCTTTGGTAGATGACATCAGGTG 792
QY 528 alleGlyGluAlaThrGlyAsnLeuGlyValHisTrpGlnLeuLysGluAspSerAsnPr 548
Db 791 TCAAGGTCAA-----CATGAGGTGTGACTCAGTTT----- 761
QY 548 oValPheAspLysThrLysAsnArgValTyrThrTyrArgAspGly-----AsnAs 566
Db 760 -ATCCTTGCAAGAGAT-----GTGACCTATCAAGATGGGAAATCAATCA 717
QY 566 nLeuMetIleGlnSer 571
Db 716 GTTGAGACTATGGAGT 701

RESULT 3
US-09-070-927A-165/c
; Sequence 165, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
; Steven Barash
; Patrick J. Dillon
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 982
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB369
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
```

```
; INFORMATION FOR SEQ ID NO: 165:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11427 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 165:
US-09-070-927A-165
```

```
Alignment Scores:
Pred. No.: 3.54e-07 Length: 11427
Score: 166.50 Matches: 121
Percent Similarity: 35.18% Conservative: 89
Best Local Similarity: 20.27% Mismatches: 215
Query Match: 4.77% Indels: 174
DB: 9 Gaps: 26

US-09-802-285A-2 (1-659) x US-09-070-927A-165 (1-11427)
QY 120 LysAspIleAsnTrpGlnMetTrpProValLysAspAsnGluValArgTrpGlnLeuHis 139
Db 7871 AAAGAGTATCGGTGGAAATCGGTATCCTCGTATGATGATCCAGATGGCTCTTTATGTTGAGT 7812
QY 140 ArgValLysTrpTrpGlnAlaMetAlaLeuValTyrHisAlaThrGlyAspGluLysTyr 159
Db 7811 CGTCAAGCTTCTCTGTAGATCTGGCACAAGCATATGACATTCTAAAAAGAACGTTAC 7752
QY 160 AlaArgGluTrpValTyrGlnTyrSerAspTrpAlaArgLysAsnProLeuGlyLeuSer 179
Db 7751 TTACAGAAATGGCACAGCTTGCTTATTGATTTTATT-----AACGATGAGGTGAGCCA 7698
QY 180 GlnAspAsnAspLysPheValTrpArgProLeuGluValSerAspArgValGlnSerLeu 199
Db 7697 AATTCGACGAATAGGAGTGTTCGCGCTCGGTAGATGTTGGGATTCGAGTAACAAAC-- 7641
QY 200 ProProThrPheSerLeuPheValAsnSerProAlaPheThrProAlaPheLeuMetGlu 219
Db 7640 -----TGGATGAAGAGCTTGACGTATATTCCATCGCTGATTTCAGA 7599
QY 220 PheLeu-----AsnSerTyrHisGlnGlnAlaAspTyrLeuSer 232
Db 7598 CTATTAGGAATTCATGATGTTTGAACAACGCTTCGTGATCCATCTCGGACTATTTCGAG 7539
QY 233 ThrHisTyrAlaGluGlnGlyAsnHisArgLeuPheGlu----- 245
Db 7538 CGGTCTCTATATCGATAAA-----TACAGCTCAGTAATTGGGGTGTGTTGGCAATTGGT 7485
QY 246 -----AlaGlnArgAsnLeuPheAlaGlyValSerPheProGluPheLysAspSerPro 263
Db 7484 GGAATGCGACTATTGATTATT-----CTTCCAGAACTGGTGACCACT-- 7440
QY 264 ArgTrpArgGlnThrGlyLysSerValLeuAsnThrGluLysLysGlnValTyrAla 283
Db 7439 AAACAGAGGGATCTAATATGTTCTCGTTTAGCTGGAACAGCTTGATCTACAAATCTATTCA 7380
QY 284 AspGlyMetGlnPheGluLeuSerProIleTyrHisValAlaAlaIleAspIlePheLeu 303
Db 7379 GATGGAATTCATTGGGAGCAGACCCCGCTGTAC----- 7347
QY 304 LysAlaTyrGlySerAlaLysArgValAsnLeuGluLysGluPheProGlnSerTyrVal 323
Db 7346 -----CAGCAGAAAGTTTGTGATGATCATCTGCG 7320
QY 324 GlnThrValGluAsnMetIleMetAlaLeuIleSerIleSerLeuPro----- 339
Db 7319 TATCTATTGCG-----ATTCTGAATATCTTGAAGTCCAGTTACCATAGATCTTCGC 7266
QY 340 ---AspTyrAsnThrProMetPheGlyAspSerTrpIleThrAspLys----- 354
Db 7265 ATGAAATTAATAAACACCTATTCTCTACCCATTATTGGCGGATAACCAAGATATCCTA 7206
QY 355 -----AsnPhe----- 356
```

Db 7205 AATCCGATCAATGATGATGATCATGTCACACTTTTCATTACGTATATGATATATCGCAAA 7146  
QY 357 -----ArgMetAlaGlnPheAlaSerTyrAla 365  
Db 7145 TTAGGTTTCATATTGGAACCTCCATGACTCGGATATGCGAAGGCTT-----TGGACG 7092  
QY 366 -----ArgValPheProAlaAsnGlnAlaLysTyrPheAla--- 378  
Db 7091 GGGGATCTTTTGAAGAAAGATCTGGGAAC-AATGAAGCCAAAGAACTTTTCGTGG 7033  
QY 379 -----ThrAspGlyLysGlnGlyLysAlaProAsnPheLeuSerLysAlaLeu 394  
Db 7032 CGAATCAAGTGGCGTGTACAAAGCAGA-----GGATATCTA----- 6991  
QY 395 SerAsnAlaGlyPheTyrThrPhe---ArgSerGlyTyrAspLysAsnAlaThrValMet 413  
Db 6990 -----TTTTCACITTTTAAACGGTCTGCATGG----- 6964  
QY 414 ValLeuLysAlaSerProGlyGlu-PheHisAlaGlnProAspAsnGlyThrPheGln 433  
Db 6963 -----GAGCGCACATGCTCATGCTTCTACAGGTGGATTTTAC 6928  
QY 433 uLeuPheIleLysGlyArgAsnPheThrProAspAlaGlyValPheValTyrSerGlyAs 453  
Db 6927 ACTCAACTACAAGGGGATGACTTATTTTCCGATAGTGTCTGTTACAGCTATGTCACAA 6869  
QY 453 pGluAlaIleMetLysLeuArgAsnTyrTyrArgGlnThrArgIleHisSerThrLeuTh 473  
Db 6867 ATCAGAGCGACTTCAGCTA-----AAAGAGTGGCGTTCGCACAAATACGATGTT 6820  
QY 473 rLeu---AspAsnGlnAsnMetValIleThr-----LysAlaArgGln 486  
Db 6819 TATCGCAGAAATCCCACTTCTAGTTTCCGATACGTGGGGTATATGACAAATTTACCGAC 6760  
QY 486 nasLysTyrGluThrGlyAsnAsnLeuAspValLeuThrTyrThrAsnProSerTyrPr 506  
Db 6759 ACCCTTATTCAGCAAAATAAAGAACTTCTCTCGGTTTTTTTGCAGATGCGGCTGGCT 6700  
QY 506 oAsnLeuAspHisGlnArgSerValPhePheIleAsnLysLysTyrPheLeu----- 523  
Db 6699 GGATANGCGGATCAGATCCATCGATTTTTCAGCGCAGCTTCATCTATTAAAGTCGAT 6640  
QY 524 -----ValIleAspArgAlaIleGlyGlu-----AlaThr 533  
Db 6639 CAACCTCCGTAGTATTATTTGATAGCTTTTCAGGACAGAAAGACTGAAATTTACAGATAC 6580  
QY 533 rGlyAsnLeuGlyValHisTyrGlnLysGluAspSerAsnProValPheAspLysTh 553  
Db 6579 CTATATTTGGCACCGTCGATAAATTTGTCAAAGAGCGCATCGGTTGCTTTAACTAC 6520  
QY 553 rLysAsnArgValTyrThrTyrArgAspGlyAsnAsnLeuMetIleGlnSerLeuAs 573  
Db 6519 GAACACAGCATAGTACACACTCTTTTCTGCTGAGGC----- 6483  
QY 573 nAlaAspArgThrSerLeuAsnGluGluGluLysValSerTyrValTyrAsnLysGln 593  
Db 6482 -----CAAACACAGCAATCATAGTGCACAAAGGTTTACAAAGCGG-AAAAGAAATTCAG 6436  
QY 593 uLeuLysArgPro-----AlaPheValPheGluLysProLysLysAsnAlaGln 609  
Db 6435 GAATGAGCCACCGCTTATCAACAAGTTTGTTCACAAACGGG-AAAAGAAATTCAG 6377  
QY 609 yThrGlnAsnPheValSerIleValTyrProTyrAspGlyGlnLysAla 625  
Db 6376 CAACAGTCATTCTCCGTTAGAGGATATCCAAATTCACGCAATAAAGT 6328

## RESULT 4

US-10-398-221-2467  
; Sequence 2467, Application US/10398221  
; Publication No. US20040018514A1  
; GENERAL INFORMATION:  
; APPLICANT: KUNST, Frederik  
; APPLICANT: GLASER, Philippe

; TITLE OF INVENTION: Listeria innocua, genome and applications  
; FILE REFERENCE: 344 702 - US  
; CURRENT APPLICATION NUMBER: US/10/398,221  
; PRIOR FILING DATE: 2003-03-27  
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061  
; PRIOR FILING DATE: 2001-10-04  
; PRIOR APPLICATION NUMBER: FR 00/12 697  
; PRIOR FILING DATE: 2000-10-04  
; NUMBER OF SEQ ID NOS: 4025  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2467  
; LENGTH: 2295  
; TYPE: DNA  
; ORGANISM: Listeria innocua  
US-10-398-221-2467

## Alignment Scores:

Pred. No.:	4,25e-05	Length:	2295
Score:	139.00	Matches:	148
Percent Similarity:	31.94%	Conservative:	97
Best Local Similarity:	19.30%	Mismatches:	280
Query Match:	3.98%	Indels:	242
DB:	16	Gaps:	37

US-09-802-285A-2 (1-659) x US-10-398-221-2467 (1-2295)

QY 7 LysArgIleIle-----ValPheAlaValIle-----AlaLeuSerSerGly 20  
Db 4 AAAAGAATTATATCTGTATTCTGCAATTTTGTCTATTGTCTAGGGATTCTTTAGCCCATTC 63  
QY 21 AsnIleLeuAlaGlnSerSerSerIleThrArgLysAspPheAspHisIleAsnLeuGlu 40  
Db 64 ACAATTAAGCAAGTGCAGCAACCAATAACAATAAATTTGAAAGAGCTT----- 114  
QY 41 TyrSerGlyLeuGluLysValAsnLysAlaValAlaAlaGlyAsnTyrAspAlaAla 60  
Db 115 -----TATCGAAGACTAA 129  
QY 61 LysAlaLeu---LeuAlaTyrTyrArgGluLysSerLysAlaArgGluProAspPheSer 79  
Db 130 AAACATTTAGGTAAACCATATTCCGACAGCAGGACAGAGAGAGAGGCGCAATTTACTTTC 189  
QY 80 AsnAla-----GluLysProAlaAspIleArgGlnPro--- 90  
Db 190 GACTGCTCAGGTACACACAATATGTTATGAAAGGTAAACAGAGATGAAGATTCTTAAT 249  
QY 91 -----IleAspLysValThrArgGluMetAlaAspLysAla 102  
Db 250 ACTTCAGCTCTCAATATTTCAGTCTGTATTAAGTGAAAAATGGTAATCAAAAACTGGT 309  
QY 103 LeuValHisGlnPheGlnProHisLysGlyTyrGlyTyrPheAspTyrGlyLysAspIle 122  
Db 310 GATTTAGTATATTTCAAAGGTCATGTAGGAATA---TACATAGGTAATGGTAAAAATGATT 366  
QY 123 AsnTyrGlnMetTyrProValLysAspAsnGluValArgTyrGlnLeuHisArgValLys 142  
Db 367 AATGCTCAAATGATGGAGTCAAAATAGACAACATT-----AATAGTAGT 411  
QY 143 TrpTyrGlnAlaMetAlaLeuValTyr-----HisAlaThrGlyAspGluLysTyrAla 160  
Db 412 TATTCGCAGAGTATTTTGTGGATACGGAAGATTTTCAATTTCTCCGAGAAAAAGGA 471  
QY 161 ArgGluTyrValTyrGlnTyrSerAsp-----Trp----- 170  
Db 472 TCTAAATCCGCTATGCTGTGTGCAGATTTAAATTTACGATCTAGCAATAACTGGGATAGT 531  
QY 171 -----AlaArgLysAsnProLeuGly-----LeuSerGlnAspAsnAspLys 184  
Db 532 TCAGTAGCTGGAAAAAGTTCCCAACGCGCTAAAGATATCTATTGACTTAGACAGTATAG 591  
QY 185 PheValTyrArgProLeuGluValSerAspArgValGlnSerLeuProThrPheSer 204  
Db 592 AATGGTTGGTGTAGTAACTTTATATAATACTAAAGGTTATATGTTTAAATACAATAAT 651

Qy	205	LeuPheValAsnSerProAlaPheThrProAlaPhe	----- 216
Db	652	TATTTCTCCGATACACCTGTGATAAAACATACTATGCCAAGGATAATATAAACCTTAAGA	711
Qy	216	-----	216
Db	712	ACGAAGCAACATGGGATAGCGATTAGCTCAAAGAAGTACAAAAGGTGAAAAAGTTACT	771
Qy	217	-----LeuMetGluPheLeuAsnSerTyrHisGlnGlnAlaAspTyrLeuSerThrHis	234
Db	772	GTGAACCTTAAACTAAATGTGATGCTTGCTTAAGTAAACATACCGCGGAAAAACAGGC	831
Qy	235	TyrAlaGluGlnGlyAsnHisArgLeuPheGluAlaGlnArgAsnLeu-----PheAla	252
Db	832	TATATGATTTTAAATAATAATTACTTAGTAAAAATCCCTTGAATATGTAACACGTATTAT	891
Qy	253	GlyValSerPheProGluPheLeuAspSerProArgTrpArgGlnThrGlyIleSerVal	272
Db	892	CGCGTTGGTACTTAAATCTCGTAGTCGCACAAACTGG---GATAGTAGCATTAAGCTTA	948
Qy	273	LeuAsnThrGlu-----IleLysLysGlnValTyrAlaAspGlyMetGlnPheGlu	289
Db	949	GTTCGTACCAGAAGCAGACGCTGTAAAGTAGAA-----ATCGATACAAT	993
Qy	290	LeuSerProIleTyrHisValala-----	297
Db	994	AGTGGTCTCTGGTTAAAGTAACATATCAAACCAACACAGGTTACATCCCACCTAACAGAT	1053
Qy	298	-----AlaIleAspIlePheLeuLysAlaTyrGlySerAlaLysArgValAsnLeu	314
Db	1054	GATTATTATCCAAAACVACTGGTCTTAAACAGTATTACGCTAAGATAATTTAAACCTA	1113
Qy	315	GluLysGlu-----PheProGlnSerTyrValGlnThrValGlu-----AsnMetIle	330
Db	1114	AGAACAAAAGCAACTGGGATAGCGACGTCTCTCAAAGGTACAAAAGGTGAAAAAGTA	1173
Qy	331	MetAlaLeuIleSerIleSerLeuProAspTyrAsnThrProMetPheGly-----	347
Db	1174	ACAGTTAATTTAAAAACAAGTGAATGGTTGGTATCAAGTAACTTATGGTGGTAAAAAA	1233
Qy	348	-----AspSerTirPileThrAspLysAsnPheArgMetAlaGlnPhe	361
Db	1234	GGTTACATGATCTTAATATGATPACTATTTAGTCGAAAAGCGGCTTAATATGAAAACCTAT	1293
Qy	362	-----AlaSerTrp-----AlaArg	366
Db	1294	TATCGGTGAGTAGCTTAAATTTACGTAGTCAGSCGAATGGGATAGTAGCATTAAGTCAA	1353
Qy	367	ValPheProAlaIleGlnAlaIleLys-----TyrPhe	377
Db	1354	GTGGTCCGAAAGGTAGAGCTGTCAAAGTCCGAATGGACACGAATGTTGCAATTTGGTTTT	1413
Qy	378	AlaThr-----AspGlyLysGlnGlyLysAlaPro-----AsnPheLeuSerLys	392
Db	1414	AAAGTAACCTATGACAAATAAACCGGTATATGCCACTTAATGATTATTATTATCTGAA	1473
Qy	393	AlaLeuSerAsnAlaGlyPheTyrThr-----PheArgSerGly	405
Db	1474	ACCGCTGTGTAAAAACCTATTATGCAAAAGATAATTTAAACTTACGTACGCAAGCAAAA	1533
Qy	406	TrpAspLysAsnAlaThrValMetValLeuLysAlaSerProProGlyGluPheHisAla	425
Db	1534	TGGGATAGTGAAGTTACTCAAAGAAGTGAAGAAAGGTGAAAAAGTAAACAGTCAATTCGAAA	1593
Qy	426	GlnProAspAsnGlyThrPheGluLeuPheIleLysGlyArgAsnPheThrProAspAla	445
Db	1594	ACRAGTATCGATGCTGGTATGAAGTA-----	1620
Qy	446	GlyValPheValTyrSerGlyAspGluAlaIleMetLysLeuArgAsnTrpTyrArgGln	465
Db	1621	-----ACATACGGTGTGAAGAAAGGCTATATGATTTTAAATAATAATATAT-----	1665

Qy	466	ThrArgIleHisSerThrLeuThrLeuAspAsnGlnAsnMetVal-----IleThrLys	483
Db	1666	---CTTGAGCAGAACCACTAGATTTTGAACAACATAATTACGCAGTTAAACATTAAATTTA	1722
Qy	484	AlaArgGlnAsnLysTrpGluThrGly-----	492
Db	1723	CGTAGTGAATCTAAATGGACAGCAGCATAAGCCAAGTGTACTGAAGCGGCTAAAGTA	1782
Qy	493	-----AsnAsnLeuAspValLeuThrTyThrAsnProSer	504
Db	1783	AAAGTTGAAATGAACACACAGGTGATGGTAAATGGTACAAAGTAACCTATCAAATAAAAAACA	1842
Qy	505	-----TyrProAsnLeuAspHisGlnArgSerValLeuPheIleasnLysLysTyrPhe	522
Db	1843	GGTTATATGCGCTAAATGAITTTATATTTANCTGAACCTGCTGGTTAAAACCTATTAT	1902
Qy	523	LeuValIleAspArgAlaIleGlyGluAlaThrGlyAsnLeuGlyValHisTrpGlnLeu	542
Db	1903	-----GCAAAAGATTAATTTAAACTTTACGTAGCGAAGCA	1935
Qy	543	LysGluAspSerAsnProValPheAspLysThrLysAsnArgValTyrThrThrTyrArg	562
Db	1936	AAATGGGATAGT-----GAAATTTCTCAAGTAGTAGAA	1968
Qy	563	AspGlyAsnAsnLeuMetIleGlnSerLeuAsnAlaAspArgThrSerLeuAsn-----	580
Db	1969	AAAGGTGAAAAGTAACATATTAATTCG-----AAAAACAGTATAAATGTTGGTGG	2016
Qy	581	-----GluGluGluGlyLysValSerTyrValTyrAsnLysGluLeuLysArgPro	597
Db	2017	CATGAAGTAACTTACGGTGGTAAAAAGGTTATATG-----ATTTTAAGTGAT	2064
Qy	598	AlaPheValPheGluLysPro	604
Db	2065	AACATATTAGTCGAAAAACCA	2085
 RESULT 5 US-10-398-221-2058/c ; Sequence 2058, Application US/10398221 ; Publication No. US20040018514AI ; GENERAL INFORMATION: ; APPLICANT: KUNST, Frederik ; APPLICANT: GLASER, Philippe ; TITLE OF INVENTION: Listeria innocua, genome and applications			

## RESULT 5

```

US-10-398-221-2058/c
; Sequence 2058, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2058
; LENGTH: 3011208
; TYPE: DNA
; ORGANISM: Listeria innocua
US-10-398-221-2058

```

Alignment Scores:		
Pred. No.:	3.82	Length: 3011208
Score:	139.00	Matches: 148
Percent Similarity:	31.94%	Conservative: 97
Best Local Similarity:	19.30%	Mismatches: 280
Query Match:	3.98%	Indels: 242
DB:	16	Gaps: 37

US-09-802-285A-2 (1-659) x US-10-398-221-2058 (1-3011208)

```

Qy      7 LysArgIleIle-----ValPheAlaValIle-----AlaLeuSerSerGly  20
Db      2117803 AAAAGAAATTAATCTGTATTGGCAATTTTCTATTGCTACTGGGATTGTTAGCCCATTC  2117744

```



```
QY 21 AsnIleuAlaGlnSerSerSerIleThrArgLysAspPheAspHisIleAsnLeuGlu 40
||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:|
Db 2117743 ACAATTAAAGCAGTCAGCAACCAATACAAATAAAACATTTGAAGAGCTT----- 2117693

QY 41 TyrSerGlyLeuGluLysValAsnLysAlaValAlaGlyAsnTyrAspAspAla 60
||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:|
Db 2117692 -----TATGCAGAGCTAAA 2117678

QY 61 LysAlaLeu-----LeuAlaTyrArgGluLysSerLysAlaArgGluProAspPheSer 79
||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:|
Db 2117677 AAAACATTTAGTAAACCATATTCGACAGACAGGACAGAGAGGCGCAATTTACTTTC 2117618

QY 80 AsnAla-----GluLysProAlaAspIleArgGlnPro--- 90
||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:|
Db 2117617 GACTGCTCAGGTATACACAAATATGTTTATGAAAAAGTAAACAGAGTAAGAAATTCCTAAT 2117558

QY 91 -----IleAspLysValThrArgGluMetAlaAspLysAla 102
||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:|
Db 2117557 ACTTCAGCTCCTCAATATCAGCTGCTGATTAAGTAAAAAATGGTAATCAAAAACCTGGT 2117498

QY 103 LeuValHisGlnPheGlnProHisLysGlyTyrGlyTyrPheAspTyrGlyLysAspIle 122
||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:|
Db 2117497 GATTTAGTATATTTCAAAGGTCATGTAGGAATA--TACATAGGTAATGGTAAAAATGATT 2117441

QY 123 AsnTrpGlnMetTrpProValLysAspAsnGluValArgTrpGlnLeuHisArgValLys 142
||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:|
Db 2117440 AATGCTCAAAATGATGGAGTCAAAATAGACAACTT-----AATAGTAGT 2117396

QY 143 TrpTrpGlnAlaMetAlaLeuValTyr-----HisAlaThrGlyAspGluLysTyrAla 160
||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:|
Db 2117395 TATTGGCAGAGATTTTTTGGTGGATCGGAGATTTTCAATTTCTCCGAGAAAAAGGA 2117336

QY 161 ArgGluTrpValTyrGlnTyrSerAsp-----Trp----- 170
||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:|
Db 2117335 TCTAAATCCGCTATGCTGTGTGCAGATTAAATTTACGATCTAGCAATAACTGGGATAGT 2117276

QY 171 -----AlaArgLysAsnProLeuGly-----LeuSerGlnAspAsnAspLys 184
||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:|
Db 2117275 TCAGTAGCTGGAAGAAAGTTCCCAAGGCGCTAAAGTATCTATTGACTTAGACAGTATAG 2117216

QY 185 PheValTrpArgProLeuGluValSerAspArgValGlnSerLeuProProThrPheSer 204
||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:|
Db 2117215 AATGGTTTGGTATGCGTAACCTTATTAATAATACTAAAGGTATATGTGTTAAATACAACTAAT 2117156

QY 205 LeuPheValAsnSerProAlaPheThrProAlaPhe----- 216
||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:|
Db 2117155 TATTCTCCGATACACCTGTGATAAAACATACTATGCCAAGGATAATATATAACTTAAGA 2117096

QY 216 ----- 216
||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:|
Db 2117095 ACGAAAGCAACATGGGATAGCGATGTAGCTCAAAAAGTACAAAAAGGTGAAAAAGTTACT 2117036

QY 217 -----LeuMetGluPheLeuAsnSerTyrHisGlnGlnAlaAspTyrLeuSerThrHis 234
||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:|
Db 2117035 GTGAACCTTAAACATAAAGTGTGAATGGTTGGTATCAAGTACATACGCGCGGAAAAACAGGC 2116976

QY 235 TyrAlaGluGlnGlyAsnHisArgLeuPheGluAlaGlnArgAsnLeu-----PheAla 252
||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:|
Db 2116975 TATATGATTTTAAATAATAATTAATTAATCTAGTAGAAAAATCCCTGGAATATGGAAACGTATTAT 2116916

QY 253 GlyValSerPheProGluPheLysAspSerProArgTrpArgGlnThrGlyIleSerVal 272
||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:|
Db 2116915 GCGGTTGGTACCTTAAATCTGCGTAGTCAGCAAACTGG---GATAGTAGCATTAGCTTA 2116859

QY 273 LeuAsnThrGlu-----IleLysLysGlnValTyrAlaAspGlyMetGlnPheGlu 289
||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:|
Db 2116858 GTTGTACCGACAGGCGAGCTGTAAAGTAGAA-----ATGATACAAAT 2116814

QY 290 LeuSerProIleTyrHisValAla----- 297
||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:|
Db 2116813 AGTGGTCCCTTGGTATAAAGTAAATATATCAATCAAAACCAACAGGTTTACATCCCACTAACAGAT 2116754

QY 298 -----AlaIleAspIlePheLeuLysAlaTyrGlySerAlaLysArgValAsnLeu 314
||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:| ||| ||| :|:|:|:|:|
```

Db 2115790 CATGAAGTAACCTTACGGTGTGTAAGAAAGGTTATATG-----ATTTTAAGTGAT 2115743

QY 598 AlapheValPheGluLysPro 604

Db 2115742 AACTATTTCGTAAGAAACCA 2115722

## RESULT 6

US-10-282-122A-34562

; Sequence 34562, Application US/10282122A

; Publication No. US20040029129A1

## GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith

; APPLICANT: Wall, Daniel

; APPLICANT: Trawick, John

; APPLICANT: Carr, Grant

; APPLICANT: Yamamoto, Robert

; APPLICANT: Forsyth, R.

; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; PRIORITY FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 34562

; LENGTH: 2088

; TYPE: DNA

; ORGANISM: Staphylococcus epidermidis

US-10-282-122A-34562

## Alignment Scores:

Pred. No.:	0.000207	Length:	2088
Score:	132.50	Matches:	140
Percent Similarity:	34.59%	Conservative:	108
Best Local Similarity:	19.53%	Mismatches:	260
Query Match:	3.79%	Indels:	210
DB:	13	Gaps:	30

US-09-802-285A-2 (1-659) x US-10-282-122A-34562 (1-2088)

QY 9 IleIleValPheAlaValIleAlaLeuSerSerGlyAsnIle---LeuAlaGlnSerSer 27

Db 88 GTATTATCTTCTGATATAGTCGGATTAGGTTATTATACAAATAGCACAAAGGATCT 147

QY 28 SerIleThrArg-----LysAspPheAspHisIleAsnLeuGluThrSerGlyLeuGlu 45

Db 148 CATTACAAACAATTATCAAAACGATGAACATAACT----- 186

QY 46 LysValAsnLysAlaValAlaAlaGlyAsn-----Tyr 56

Db 187 ---GTTAATGAATCAGTACCAGAGCGCAATACTAGATAGAAATGCAAAAGTACTAGTT 243

QY 57 AspAspAlaAlaLysAlaLeuLeuAlaTyrTyrArgGluLysSerLysAlaArgGluPro 76

Db 244 GATAATGCTTCAAGAGATGCTATTACATACACAGAACCCGTAAACATCACAAAGGAA 303

QY 77 AspPheSerAsnAlaGluLysProAlaAsp---IleArgGlnProIleAspLysValThr 95

Db 304 ATGTTAATAACTGCTAAGAAACTCACAGATTTTAATTAAATGGATACAGATAAAATACT 363

QY 96 ArgGluMetAlaAspLysAlaLeuValHisGlnPheGlnProHisLysGlyTyrGlyTyr 115

Db 364 GAGAGA-----GATAAA----- 375

QY 116 PheAspTyrGlyLysAspIleAsnTyrGlnMetTyrPro-----ValLysAspAsn 132

Db 376 -----AAGGATTTTGGATTCAAAATGTATCCGTCATCTGCTAAAAAGTT-AAT 422

QY 133 GluValArgTyrGlnLeuHisArgValLysTyrTyrGln----- 145

Db 423 GAGAAAAGA-----ACAATTAAATGTTAGAGGATGCGAGTATTTTCAACAGCAATTTGA 476

QY 146 ----- 147

Db 477 TACCCAACTTAGAGATAAAATAGGAAAAACAATTAACACAGTTAACTAAAAAGATTT 536

QY 148 AlaLeuValTyrHisAlaThrGlyAspGluLysTyrAla----- 160

Db 537 GCAAGTTTACGCAATTTATCGGGAATGACGCTGGTCACTCTAGATCCTCAACAAT 596

QY 161 -----ArgGlu----- 162

Db 597 TAAAAATGAAGACGTAAAGCGAGAAAGAAATATGCGCGGTATCAACAGAGCTTTCTAAAT 656

QY 163 ---TipValTyrGlnTyrSer-AspTyrAlaArgLysAsnProLeuGlyLeuSerGlnAs 181

Db 657 ACCTGGTGAATAATACATACAAATGATGGGATGGAATAACCCATACCGT-----GA 707

QY 181 pAsnAspLysPheValTyrArgProLeuGluValSerAspArgValGlnSerLeuProPr 201

Db 708 TACTTTAAGAGGCATATTTGGA-----GATGTGTGCGACTTCGACTGAAGGTATACCTAA 761

QY 201 oThrPheSerLeuPheValAsnSerProAlaPheThrProAlaPheLeuMetGluPheLe 221

Db 762 AGAATTAACCT-----GAACAATATTT 782

QY 221 uAsnSerTyrHisGlnGlnAlaAspTyrLeuSerThrHisTyrAlaGluGlnGlyAsnHi 241

Db 783 ATCAAAAGGTTATTACGAAATGATCGGGTCGTAATCTTTATCTTGAATATCAATACGA 842

QY 241 sArgLeuPheGluAlaGlnArgAsnLeuPheAlaGlyValSerPheProGluPheLysAs 261

Db 843 AGATGTTCTTAAAGGCACGAGAAACAATG-----AAATATACAAC 884

QY 261 pSerProArgTyrArgGlnThrGlyTyrSerValLeuAsnThrGluIleLysGlnVa 281

Db 885 TGATAAATCTGGAAGAGTAATAAGTTCAAGAGTACTTAATCTCTGGCTCAAGA----- 936

QY 281 lTyrAlaAspGlyMetGlnPheGluLeuSerProIleTyrHisValAlaAlaIleAspI 301

Db 937 -----GGTCATGATTTACAATTACT-----ATAGATAT 965

QY 301 ePheLeuLysAlaTyrGlySerAlaLysArgValAsn-----LeuGluLysGluPhePr 319

Db 966 TGATTTTACAG-----AAAAAGTAGAATCTTTATTAGAAAAACAATTTTC 1010

QY 319 oGlnSerTyrValGlnThrValGluAsnMetIleMetAlaLeuIleSerIleSerLeuPr 339

Db 1011 TAAATTACGTACAGGTGCTAAGGATATGCAATGCGTTAATGGTTGCTCCAAATTC 1070

```
Qy 339 oAspTyrAsnThrProMetPheGlyAspSerTrpIleThrAspLysAsnPheArgMetAl 359
|
Db 1071 TAAAAATGGAGACATTTCTCGCTATTTCAGGAAAGCAAAATTGATAAGCAAGGTAAACTCAA 1130
|
Qy 359 aGlnPhe-----AlaSerTrpAlaArgValPheProAlaAsnGlnAlaIleLys---Ty 376
|
Db 1131 AGATTATGATTCGGCACTTTACAGCTCAATACACAGTAGGTCTTCAGTAAAGGAGG 1190
|
Qy 376 rPheAlaThrAspGlyLysGlnGlyLysAlaProAsn-----PheLeuSerLys 392
|
Db 1191 AACATTATTAGCTGACCAACAAATAAAGCTATTAACTGTCGAGAAACATGCTAGATGA 1250
|
Qy 392 sAlaLeuSerAsnAlaGlyPheTyrThrPheArgSerGlyTrpAspLysAsnAla----- 410
|
Db 1251 GCCATTAAATATCCAAAGGTGGTTTAACTAAGCGTTCTTATTATTAATAAAATGCTCATGT 1310
|
Qy 411 -----ThrVal 412
|
Db 1311 ATCTATCGATGATAAACAGCACCTTATGCTATCATCAAAAGTATACATGTTTAAACCGC 1370
|
Qy 412 lMetValLeuLysAlaSerProProGlyGluPheHisAlaGlnProAspAsnGlyThrPh 432
|
Db 1371 ACTTAAATTAGCAGTGACCCCTATCTACTTCAGGTATGCTATTACCTAATAAT----- 1422
|
Qy 432 eGluLeuPheIleLysGlyArgAsnPheThrProAspAlaGlyValPheValTyrSerGl 452
|
Db 1423 -----ATAGCAGATGCTGCT----- 1437
|
Qy 452 yAspGluAlaIleMetLysLeuArgAsnTrpTyrArgGlnThrArgIleHisSerThrLe 472
|
Db 1438 -----CGTAAATTCGGTAAAGGGTTAAATCAAGTAGTCTTCGCTTAAAC 1484
|
Qy 472 uThrLeuAspAsnGlnAsnMetValIleThrLysAlaArgGlnAsnLysTrpGluThrGl 492
|
Db 1485 AGGTATTGACTTACCGAAGCAAGCCAGCCAAATAGAACCTTAACTTAATATCTCTGG 1544
|
Qy 492 yAsnAsnLeuAspValLeuThrTyrThrAsnProSerTyrProAsnLeuAspHisGlnAr 512
|
Db 1545 TAACATTATTAGCTTAGCTATTGGACAATACACACATATACACCACCTTCAGTTCTCCCA 1604
|
Qy 512 gSerValLeuPheIleAsnLysLysTyrPheLeuVal-----IleAspArgAlaIle 529
|
Db 1605 ATACGTATCAACTATGCTAATGCTATGCTATAGATTCACACCACATATGGATTGCTAT 1664
|
Qy 529 eGlyGluAlaThrGlyAsnLeuGlyValHisTrpGlnLeuLysGluAspSerAsnProVa 549
|
Db 1665 TTATGAACTCTACTAAT-----AAAGATGAAACTGGTCCATT 1700
|
Qy 549 lPheAspLysThrLysAsnArgValTyrThrTyrArgAspGlyAsnAsnLeuMetIl 569
|
Db 1701 AAACGTAAATTAAGGTAAATGTTTTTAAATAGGTAAATAACTCAAAATGACGAAATTA 1760
|
Qy 569 eGlnSerLeuAsnAlaAspArgThrSerLeuAsnGluGluGly-LysValSerTyrV 589
|
Db 1761 AGAAGTTCAAGAGGCTTCAAAATGGCTTCAATGAAAGCAAGGTACAGGTTATGCTAG 1820
|
Qy 589 al-----TyrAsnLysGluLeuLysArgProAlaPheValPheGluLysP 604
|
Db 1821 TTTTGAATAATCTGTAGTACTCTTCAGCTGGTAAACAGCAAGTCTGTAAGTTTTCAGA 1880
|
Qy 604 roLysLysAsnAlaGlyThrGlnAsnPheValSerIleValTyrPro 619
|
Db 1881 CGGAGAACCTAGATGTTAACTCAACAT---ATATCGGTATGACCCCG 1924
```

## RESULT 7

US-10-282-122A-35377

; Sequence 35377, Application US/10282122A

; Publication No. US20040029129A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Liangsu

; APPLICANT: Zamudio, Carlos

; APPLICANT: Malone, Cheryl

; APPLICANT: Haselbeck, Robert

```
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35377
; LENGTH: 2043
; TYPE: DNA
; ORGANISM: Staphylococcus haemolyticus
US-10-282-122A-35377
```

## Alignment Scores:

Pred. No.:	0.000342	Length:	2043
Score:	130.50	Matches:	134
Percent Similarity:	33.08%	Conservative:	88
Best Local Similarity:	19.97%	Mismatches:	256
Query Match:	3.73%	Indels:	193
DB:	13	Gaps:	27

US-09-802-285A-2 (1-659) x US-10-282-122A-35377 (1-2043)

```
Qy 9 IleIleValPheAlaValIleAlaLeuSerSerGlyAsnIle---LeuAlaGlnSerSer 27
|
Db 88 GTAGTTATATTGGAAATTATAGTTTAAAGTTAGGATATTACAAATTCGCAAGGCTCC 147
|
Qy 28 SerIleThrArgLysAspPheAspHisIleAsnLeuGluTyrSerGlyLeuGluLys--- 46
|
Db 148 -----CAATACAGCCAACTAGTTAAAAAC 171
|
Qy 47 -----ValAsnLysAlaValAlaAlaGlyAsn----- 55
|
Db 172 GATGNAACATACATGTTAAATGAATCTCTTCTAGGGGAAGAAATTTCTGATAGGAATGGA 231
|
Qy 56 -----TyrAspAspAlaAlaLysAlaLeuLeuAlaTyrTyrArgGluLysSerLys 72
|
Db 232 AAAAGTATTAGTAGATAATGCTTCTTAAACTAAACGATTATACACAAAGATCTAGGAAACA 291
|
Qy 73 AlaArgGluProAspPheSerAsnAlaGluLysProAlaAsp---IleArgGlnProIle 91
|
Db 292 AGTCAAAAGATATGCTCGATGATCTGCAAAAAAAATTTGTCATCTCTCTATTACTATGAAACT 351
|
Qy 92 AspLysValThrArg----- 96
|
```

352 GACAAAATTACTGAAAGAGATAAACAAGATTTTGGATTTCAGAAACACCAAGATGAGTA 411  
QY 97 -----GluMetAlaAspLysAlaLeuValHisGlnPheGln 108  
Db 412 GACAAATTAATGAAAAAGAACTTCAATGTTAAATGAAGGTAGTATTACCAAGATCAA 471  
QY 109 ProHisLysGly----- 112  
Db 472 TATGATAAACAATATATAAAGAGTTGGAGACAAACAATAATAGTTTATCAAAAAA 531  
QY 113 -----TyrGlyTyrPheAspTyrGlyLysAspIleAsnTrpGln 125  
Db 532 GACTTACGAGTTTACCAATTTATCGTGAATGCTGCTGTTCTACCAATGACCCGCA 591  
QY 126 MetTrpProValLysAspAsnGluValArgTrpGlnLeuHisArgValLysTrpTrpGln 145  
Db 592 -----ACAAATTAAGATGAGATGTAAGTGAATAA----- 621  
QY 146 AlaMetAlaLeuValTyrHisAlaThrGlyAspGluLysTyrAlaArgGluTrpValTyr 165  
Db 622 -----CAATATGCTGCAGTGTCAACAACCTTGATGCTTCAGGTGTAAT 669  
QY 166 GlnTyrSerAspTrpAlaArgLysAsnProLeuGlyLeuSerGlnAspAsnLysPhe 185  
Db 670 ACAACTATGATGGGACCGACGCTATCTTATGGT----- 705  
QY 186 ValTrpArgProLeuGluValSerAspArgValGlnSerLeuProProThrPheSerLeu 205  
Db 706 -----GATACATTAAAGAGTATA-----TTTGGTAGT 732  
QY 206 PheValAsnSerProAlaPheThrProAlaPheLeuMetGluPheLeuAsnSerTyrHis 225  
Db 733 GTATCTACATCAAGTGAAGTATCTCTAAGAATTGACCGAA----- 774  
QY 226 GlnGlnAlaAspTyrLeuSerThrHisTyrAlaGluGlnGlyAsnHisArgLeuPheGlu 245  
Db 775 -----CAATATTAGCTAAAGGTATTCT----- 798  
QY 246 AlaGlnArgAsnLeuPheAlaGlyValSerPheProGlu-----PheLysAspSerPro 263  
Db 799 -----CGGAATGATAGAGTTGGTAAATCCTATTAGAGTATCAATATGAAGATATTTA 852  
QY 264 ArgTrpArgGlnThrGlyIle-----SerValLeuAsnThr 275  
Db 853 CGTGGTAAGAAAAAGAAATGAATATCTACTGATAAATCTGAAAAAGTCATAACTCT 912  
QY 276 Glu-----IleLysLysGlnValTyrAlaAspGlyMetGlnPheGluLeuSerProIleTyr 294  
Db 913 GAAGTCATTAATCCTCGATCTAGAGCGGATGATTACAGTTA----- 954  
QY 295 HisValAlaAlaIleAspIlePheLeuLysAlaTyrGlySerAlaLysArgValAsn--- 313  
Db 955 -----ACAATTGATATAGCTTACAG-----AAGAAAGTAGAATCT 990  
QY 314 ---LeuGluLysGluPheProGlnSerTyrValGlnThrValGluAsnMetIleMetAla 332  
Db 991 TTATTAGAAAAATCAAAATCAAGACGTTACGAGTCAAGGTCTAAGATATGATAACGCG 1050  
QY 333 LeuIleSerIleSerLeuProAspTyrAsnThrProMetPheGlyAspSerTrpIleThr 352  
Db 1051 CTTATTGTGTGCAAAATCCTAAAAATGGCGATATTTTAGCTATGCGAGAAACAAATC 1110  
QY 353 AspLysAsnPheArgMetAlaGlnPhe-----AlaSerTrpAlaArgValPheProAla 370  
Db 1111 GATAAAATGGTAAGCTAACAGATTATGATTAGGGAACCTTTACTGTCATTTGCAGTT 1170  
QY 371 AsnGlnAlaIleLys-----TyrPheAlaThrAspGlyLysGln 383  
Db 1171 GGTCTCTTCAGTAAAGAGTGGACCCCTATTAGCTGGATATCAAAATATGCGATTAAAGTT 1230  
QY 384 GlyLysAlaProAsnPheLeuSerLysAlaLeuSerAsnAlaGlyPheTyrThrPheArg 403  
Db 1231 GCGCAG-----GAAATGATAGATGAGCCACTTCATTTAAAGGTGGATTACAAAGCGC 1284

QY 404 SerGlyTrpAspLysAsnAlaThrValMetValLeuLysAlaSerProProGlyGluPhe 423  
Db 1285 TCCTACTTTAATAAATAAATGATAAAGTCAGAAAT-----AATGATAAA 1326  
QY 424 HisAlaGlnProAspAsnGlyThrPheGluLeuPheIleLysGlyArgAsnPheThrPro 443  
Db 1327 GAAGCAATTAATGACACTCTTCAACACGATATATGTTT-----AAAACGTCTTTAAA 1377  
QY 444 AspAlaGlyValPheValTyrSerGly-----AspGluAlaIle 456  
Db 1378 TTAGCCGGAGATCCTTACTATATAGTGAATGGGCTTACCTACAGATATAAGTGAAGCTGGC 1437  
QY 457 MetLysLeuArgAsnTrpTyrArgGlnThrArgIleHisSerThrLeuThrLeuAspAsn 476  
Db 1438 CAGAAATTAAGAAAAAGGGCTTAATCAAGTCGAGTGGAGTTTAAACACAGGATTGACTTG 1497  
QY 477 GlnAsnMetValIleThrLysAlaArgGlnAsnLysTrpGluThrGlyAsnLeuAsp 496  
Db 1498 CCAATTAACAACATAGTCAATTTGAACCATTAACAACAATTCAGGCAATTTATCTGAT 1557  
QY 497 ValLeuThrTyrThrAsnProSerTyrProAsnLeuAspHisGlnArgSerValLeuPhe 516  
Db 1558 TTATCAATTTGGTCAATATGACACATATAGTCCAAATTCATATACACAGTACGTTCTACA 1617  
QY 517 IleAsnLysLysTyrPheLeuVal-----IleAspArgAlaIleGlyGluAlaThr 533  
Db 1618 ATAGCTAATGATGGTTATAGAAATTCACACCACACATTTGGATTAGCAATTCATGATGTACA 1677  
QY 534 GlyAsnLeuGlyValHisTrpGlnLeuLysGluAspSerAsnProValPheAspLysThr 553  
Db 1678 AATAGT-----GATGACATAGTCCAGTTAAACAAAAAATT 1713  
QY 554 LysAsnArgValTyrThrTyrArgAspGlyAsnLeuMetIleGlnSerLeuAsn 573  
Db 1714 AAAGGCAATGCTCTTAAACAAAGTAAATTAATTCAGAAGATCAATTAAGAAAGTACAAAAA 1773  
QY 574 AlaAspArgThrSerLeuAsnGluGluGly 584  
Db 1774 GGATTTGAAATGGCATTCATCAATGAAAGATGGA 1806

## RESULT 8

US-10-369-493-45761  
; Sequence 45761, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; PRIOR FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 45761  
; LENGTH: 1599  
; TYPE: DNA  
; ORGANISM: Saccharomyces cerevisiae  
US-10-369-493-45761

Alignment Scores:  
Pred. No.: 0.00171 Length: 1599  
Score: 123.00 Matches: 109  
Percent Similarity: 32.01% Conservative: 76  
Best Local Similarity: 18.86% Mismatches: 179  
Query Match: 3.52% Indels: 214  
DB: 16 Gaps: 31

Score: 122.50 M

Percent Similarity:	32.91%	Conservative:	98
Best Local Similarity:	20.41%	Mismatches:	296
Query Match:	3.51%	Indels:	230
DB:	10	Gaps:	37
US-09-802-285A-2 (1-659) x US-09-769-736-9 (1-3753)			
QY	24	AlaGlnSerSerIleThrArgLysAspPheAsp	-----HisIleAsnLeu 39
DB	1303	AGTAAATCTCTTTAATATCAAGCGGACTTTAATCTTAACCAAGTCAITTCATATATA	1362
QY	40	GluTyrSerGlyLeuGluLysValAsnLysAlaValAlaAlaGlyAsnTyrAspAla 59	
DB	1363	TCATTATAATGGT-----AACAAATGCACGACAGCAATCTTCG---GAATTT 1407	
QY	60	AlaLysAlaLeuLeuAlaTyrTyrArgGluLysSerLysAlaArgGluProAspPheSer 79	
DB	1408	AAAGACCAACTTTATGCTTTATAGTGGAAATTTAGGTGCAGTCTTCAATCAAGATGGTTCA 1467	
QY	80	AsnAlaGluLysProAlaAspIleArgGlnPro---IleAspLysValThrArgGluMet 98	
DB	1468	AAAGTTGAA-----GCCAGCCTCTGGTCACCGAGTGTGTAGTGCATATGATTAT 1521	
QY	99	AlaAspLys-----AlaLeuValHisGlnPheGlnProHisLys 111	
DB	1522	TATGACAAAGATAATCAAAACAGGGTTGTAGCGACTACCCCTCTGTGAAAAATAATAA 1581	
QY	112	GlyTyr-----GlyTyrPhe 116	
DB	1582	GGTGTGTGGCAGCAGTACTTGATGATAATAGGTATTAAAAACTATATCTGGTTACTAT 1641	
QY	117	-----AspTyrGlyLysAspIleAsnTyrGlnMetTyrProValLysAsp 131	
DB	1642	TATCTTTACGAATAAAACAGAGTAAGGATAAGTTAAGATTTTAGATCCTTATGCAAG 1701	
QY	132	AsnGluValArgTyrGlnLeuHisArgVal-----LysTyrTyrGlnAlaMet 147	
DB	1702	TCATTAGCAGTGGGAGTAGTAAATACGTCTTAATGACGATATAAAACGGCTAAAGCAGCT 1761	
QY	147	----- 147	
DB	1762	TTTGTAATCCAAAGTCAACTGGACCTAAATAATTTAAGTTTTCCTAAAAATGCTAATTTT 1821	
QY	148	-----AlaLeuValTyrHisAlaThrGlyAspGluLysTyrAlaArgGlu 162	
DB	1822	AAAGGAAAACAGATCTGTATATAC-----GAAGCACATGTAAAGAC 1866	
QY	163	TyrValTyrGlnTyrSer---AspTyrPalaArgLysAsnProLeuGly---LeuSerGln 180	
DB	1867	TTCACTTCTGATCAATCTTTGGACGGGAAATTAATAAATCAACTTGGTACCTTTGCAGCC 1926	
QY	181	AspAsnAspLysPheValTyrArgProLeuGluValSerAspArgValGlnSerLeuPro 200	
DB	1927	TTTTTCAGAGAACTAGATTATTACAAATTAGGAGTTACACACATTCAGCTTTTACCG 1986	
QY	201	ProThrPheSerLeuPheValAsn-----SerProAlaPheThrPro 214	
DB	1987	GTATTGAGTTATTTTATGTGTTAATGAATGGAATGAGTACAGCTCAACAGCTTACACTTC 2046	
QY	214	----- 214	
DB	2047	TCAGACAATAATACAAATTTGGGCTATGACCCACAGAGCTATTTTGCTCTTTCTGGAATG 2106	
QY	215	-----AlaPheLeuMetGluPheLeuAsn 222	
DB	2107	TATTCAGAGAAACCAAGATCCATCAGCAGCTATCGCGAATTAACAAATTAATACAT 2166	
QY	223	SerTyrHisGlnGlnAla-----AspTyrLeuSerThrHisTyrAlaGlu 237	
DB	2167	GATATTCAATAACGTGGCATGGGGTTATACCTTGTCTTATAATACACACTGCAAAA 2226	
QY	238	GlnGlyAsnHisArgLeuPheGlu---AlaGlnArgAsnLeuPheAlaGlyValSerPhe 256	
DB	3118	GCA-----GGTTTAATACATTAAAGCGGTTCAACAGATGCTTTCGGAAATTGAGC 3168	
DB	2227	ACT-----TATCTCTTTTGAGGATATAGAACCTAATATTATAT-----CACTTT 2268	
QY	257	ProGluPheLysAspSerProArgTyrArgGlnThrGly----- 269	
DB	2269	ATGAATCAAGATGGTTTCCACCAAGAGAAAGTTTGGAGGGGAGCGTTTAGGAACCACTCAT 2328	
QY	270	-----IleSerValLeuAsnThrGluIleLysLys 279	
DB	2329	GCAATGAGTCGTCGTGTTTGGTTGGTTCATTAAATATCTTCAAGTGAATTTAA----- 2385	
QY	280	GlnValTyrAlaAspGlyMetGlnPheGluLeuSerProIleTyrHisValAlaAlaIle 299	
DB	2386	-----GTTGATGGTTTCCGTTTGTATGATCATGGAGATCATGATGGCTGGGATT 2436	
QY	300	AspIlePheLeuLysAlaTyrGlySerAlaLysArgValAsnLeuGluLysGluPhePro 319	
DB	2437	GAATTA-----GCTTATAAAGAGCTAAAGCTAATTAAT----- 2469	
QY	320	GlnSerTyrValGlnThrValGluAsnMetIleMetAlaLeuIleSerIleSerLeuPro 339	
DB	2470	-----CCTAATATGATT----- 2481	
QY	340	AspTyrAsnThrProMetPheGlyAspSerTyrPheThrAspLysAsnPheArgMetAla 359	
DB	2482	-----ATGATTGGTGAGGGCTGG-----AGAACAATC 2508	
QY	360	GlnPheAlaSerTrpPalaArgValPheProAlaAsnGln----- 372	
DB	2509	CAAGGCGATCAAGGTAAAGCCGTTAAACCCAGCTGACCAAGATTGGATGAAGTCAACCGAT 2568	
QY	373	AlaIleLysTyrPheAlaThrAspGlyLysGlnGlyLysAlaProAsnPheLeuSerLys 392	
DB	2569	ACAGTTGGCGCTCTTTCAGATGATTCGT-----AATAGCTGAATCT 2613	
QY	393	AlaLeuSerAsnAlaGlyPheTyrThrPheArgSerGlyTyrAspLysAsnAlaThrVal 412	
DB	2614	GGTTTTCCAAATGAAGGTACTCCAGCTTTCATCACAGGTGGCCCACAATCTTTACAAGGT 2673	
QY	413	MetVal-----LeuLysAlaSerPro-----ProGly----- 421	
DB	2674	ATTTTTAAATATCAAGCAACCACTGGGAATTTTGAACGAGATTTCGCAGAGATGTG 2733	
QY	422	---GluPheHisAlaGlnProAspAsnGlyThrPhe---GluLeuPheIleLysGlyArg 439	
DB	2734	GTGCAATATATTGCTCCACATGATACCTTACCTTCGATCATGTGATGTCAAATCAATT 2793	
QY	440	AsnPheThrProAspAlaGlyValPheValTyrSerGlyAspGluAlaIleMetLysLeu 459	
DB	2794	AATAAAGACCTTAAGTA-----GCTGAAGAGATATT----- 2826	
QY	460	ArgAsnTyrTyrArgGlnThrArgIleHisSerThrLeuThrLeuAspAsnGlnAsnMet 479	
DB	2827	-----CATGACGTCTGCTTTAGGAATGATTAATGATTTTAACTCTCAAGGACA 2877	
QY	480	ValIleThrLysAlaArgGlnAsnLysTyrGluThrGlyAsnAsnLeuAspValLeuThr 499	
DB	2878	GCATTCATTTCATTCGTGTCAGAGTATGTCGTACGAGCGTTTACTTAACCTGATTAC 2937	
QY	500	TyrThrAsnProSerTyrProAsnLeuAspHisGlnArgSerValLeuPheIleAsnLys 519	
DB	2938	ATGACAAAAAGTTTCAGATGACAAATTCCTTAATAAAGCAACACTTATTGAAGCTGTTAAA 2997	
QY	520	LysTyrPheLeuValIleAspArgAlaIleGlyGluAlaThrGlyAsnLeuGlyValHis 539	
DB	2998	GAATACCCATATTTTATTCATGATCATATGATCTTCAGATGCCATTAATCATTTTGTAT 3057	
QY	540	TrpGlnLeuLysGluAspSerAsnProValPheAspLysThrLysAsnArgValTyrThr 559	
DB	3058	TGGGAGCAGCCACAGATAATAACAAACACCAAAATTTCAACGAAAAACACAGGCTATACA 3117	
QY	560	ThrTyrArgAspGlyAsnAsnLeuMetIleGlnSerLeuAsnAla-----Asp 575	
DB	3118	GCA-----GGTTTAATACATTAAAGCGGTTCAACAGATGCTTTCGGAAATTGAGC 3168	

Qy	576	ArgThrSerLeuAsnGluGluGlyValSerTyrValTyrAsnLysGluLeuLys	595
	:		
Db	3169	AAAGCAGAAATTCATCGCTAGGTTAGCTTGATACAGAGGTAGCTCAAGGTGATATTAAA	3228
	:		
Qy	596	ArgProAlaPheValPheGluLysProLysIysAsnAlaGlyThrGlnAsnPheValSer	615
	:		
Db	3229	GAAAAAGATTGGTTATT-----GCTTACCAACAATAGATTCTAAAGCGGATATTTC	3282
	:		
Qy	616	IleValTyrProTyrAspGlyGlnLysAlaProGluLeuSerIleArgGlu-AsnLysG1	635
	:		
Db	3283	GCAGTATTGTTAATGCTGATAGTAAAGCTAGAAACGTTTTACTAGGTGAAAAATATA--	3340
	:		
Qy	635	YAsnAspPheGluLysGlyIysLysLeuAsnLeuThrLysIleAsnGlyLysGlnGlnLe	655
	:		
Db	3341	-AACACCTTTTAAAGGCGAAGTAATTGTTGATCTGATCAAGCGGGGATTAAACCAATC	3399
	:		
Qy	655	uValLeuVal	658
	:		
Db	3400	TCAACTCCTA	3409

```

RESULT 10
US-10-320-797-2302
; Sequence 2302, Application US/10320797
; Publication No. US20040014955A1
; GENERAL INFORMATION:
; APPLICANT: Eroshkin, Alexey M.
; APPLICANT: Zamudio, Carlos
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
; TITLE OF INVENTION: METHODS OF USE
; FILE REFERENCE: 10182-021-999
; CURRENT APPLICATION NUMBER: US/10/320,797
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341,261
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 3361
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2302
; LENGTH: 2541
; TYPE: DNA
; ORGANISM: Cryptococcus neoformans
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1666)..(1666)
; OTHER INFORMATION: n=g, a, t or c
US-10-320-797-2302

```

Alignment Scores:	
Pred. No.:	0.0061
Score:	121.00
Percent Similarity:	36.55%
Best Local Similarity:	20.52%
Query Match:	3.46%
DB:	16
	Length: 2541
	Matches: 151
	Conservative: 118
	Mismatches: 252
	Indels: 222
	Gaps: 37
US-09-802-285A-2 (1-859) x US-10-320-797-2302 (1-2541)	

Qy	8	ArgIleIleValPheAlaValIleAlaLeuSerGlyAsnIleLeuAlaGlnSer	27
Db	106	AGGGTGTCTGTTTCATCTGGTGAAACAACAGTCAAGGACATTCGCCCTTTGAGAAC---	162
Qy	28	SerIleThrArgLysAspPheAspHisIleAsnLeuGluTy-SerGlyLeuGluLysVal	47
Db	163	-----AGAAAATAATCCGCTTCAATTGCCTCTTCTCCAGATGCG-----	201
Qy	48	AsnLysAlaValAlaAlaGlyAsnTrpAspAlaAlaLysAlaLeuLeuAlaTrpYr	67
Db	202	AATGTAATAATT-----TCCATCGACGAAGATGGAAGACACTGTTTGGCACTTC	252
Qy	68	ArgGluLysSerLysAlaArgGluProAspPheSerAsnAlaGluLysProXlaAspIle	87
Db	253	CCTAAAGGGGACA-----	264

Qy	88	ArgGlnProIleAspLysValThrArgGluMetAlaAspLysAlaLeuValHis	-----	105
Db	265	-----		
Qy	106	GlnPheGlnProHisLysGlyTyrGlyTyrPheAspTyrGlyLysAspIleAsnTrpGln	-----	125
Db	307	TCATTTTCTCCGAGCGCAAAATATTATTCGATACGCACGGTCACATGGTG	-----CAA	360
Qy	126	MetTrpProValLysAspAsnGluValArg	-----TrpGlnLeuHisArg	140
Db	361	ATATGGAACTCCACCCATTTGGTCCGAGAGTTTCACCTTTTACCTCCATAGGAG	-----	420
Qy	141	-----	ValLysTrpTrpGlnAlaMetAlaLeuVal	150
Db	421	TATACTGGTCATCACGACGAAGTTGTAGTGTCTGCTGGTC-CAAAACCTCCAGGTATTT	-----	479
Qy	151	TyrHisAlaThrGlyAspGluLysTyrAlaArgGlnTrpValTyrGln	-----	166
Db	480	TATCATTACCTC-----GAGAGATATGACTGCGAGGCTATATACCATTAACCCACTTGA	-----	533
Qy	167	-----	-----Tyr	167
Db	534	AGGTTTCCAACTAAGCAATTTGGGGACATAGAGATGTTGTGTGGAGCGTCTTTTC	-----	593
Qy	168	Ser-AspTrpAlaArgLysAsnProLeuGlyLeuSerGlnAspAsnAspLysPheValTr	-----	187
Db	594	TCAAGAT-----GAGAAGACTATTATACACGGTTTCTCGAGACGGTCAGTGTGTGATG	-----	647
Qy	187	PAArgProLeuGlu--ValSerAsp-ArgValGlnSerLeuProProThrPheSerLeuP	-----	206
Db	648	GAAGCAAAAGAGCGCTCTGAGCGGACTCGGATGTTGAGATGACATTTCTCGAATC	-----	707
Qy	206	heValAsnSerProAlaPheThrProAlaPheLeuMetGluPheLeuAsnSerTyrHisG	-----	226
Db	708	TCCACCACCTCCACCTCGCGCAACCTTGCC-CTTGAGACGCGCTCGCTAC	-----	762
Qy	226	lnGlnAlaAspTyrLeuSerThrHisTyrAlaGlnGlnGlyAsnHisArgLeuPheGluA	-----	246
Db	763	-----	-----	794
Qy	246	laGlnArgAsnLeuPheAlaGlyValSerProGluPheLysAspSerProArgTrpA	-----	266
Db	795	-----	-----CAACCAACCTGGT	807
Qy	266	rgGlnThrGlyLysSerValLeuAsnThrGluIleLysLysGlnValTyrAlaAspGlyM	-----	286
Db	808	ACGAAGTGATTTGGCTACTCTCCACCCAA-----	-----	839
Qy	286	etGlnPheGluLeuSerProIleTyrHisValAlaIleAspIlePheLeuLysAlaT	-----	306
Db	840	-----	-----	885
Qy	306	yrGlySerAlaLysArgValAsnLeuLysGluPheProGlnSerTyrValGlnThrV	-----	326
Db	886	TGGAGATGCC-----CGAATTTACCCCGTACATCGTTG--TCGA	-----	925
Qy	326	alGluAsnMetIleMetAlaLeuIleSerLeuProAspTyrAsnThrProMetP	-----	346
Db	926	TCTCCAACGAGAGATCTCTAGTGTGGCAGTCTCCGG-----	-----	963
Qy	346	heGlyAspSerTrpIleThrAspLysAsnPheArgMetAlaGlnPheAlaSerTrp	-----	364
Db	964	--TCGGAGAGTGGTGGCATTTGGGCGCGCAAGCTCGCAGATTTATGTTTGGGAAT	-----	1021
Qy	365	-----AlaArgValPheProAlaAsnGlnAlaIleLysTyrPhe	-----	377
Db	1022	GGCAGATGAGAGTTACGTTCTTAAACAGCAAGAGTCTACTACTACGACATGAACACCTCG	-----	1081
Qy	378	-----	-----AlaThrAspGlyLysGlnGlnGlyLysAlaProAsnP	389
Db	1082	CGTTTAGTCCGATGGGCAGAAATCATCTACTGCGGTGAAGTGTGAAGTCAAG-----	-----	1137
Qy	389	heLeuSerLysAlaLeuSerAsnAlaGlyPheTyrThrPheArgSerGlyTyrAspLysA	-----	409

```
Db 1138 --TTATGGAATGCTCAGTGGCTTCGCTTTGTGACCTTCCT-----GAACACA 1186
QY 409 snAlaThrValMetValLeuLysAlaSerProProGlyGlu-----PheHisAlaGlnP 427
Db 1187 CTGCGGTATCTCCACTGTGCAATTTGCTAAGCAGGACAACTTTTATTACACAGCGTCC 1246
QY 427 roAspAsnGlyThrPheGluLeuPhe-----lleLysGlyArgAsnPheThrProAspA 445
Db 1247 TTGAC---GGTACTGTCCGCGCATACACCTCATCCGATACCGTAACCTC----- 1293
QY 445 laGlyValPheValTyrSerGlyAspGluAlaIleMetLysLeuArgAsnTyrPheArgG 465
Db 1294 -----CGACATTCACCTCTC 1309
QY 465 lnThrArgIleHis---SerThrLeuThrLeuAspAsnGlnAsnMetValIleThrLysA 484
Db 1310 CCACCCCTGCTCCAGTTCTCTGCGCTCGCGCTCGATCCTTCAGCGCATGCTGCTGCTG 1369
QY 484 laArgGlnAsnLysTrpGlu-----ThrGlyAsnAsnLeuAspV 497
Db 1370 GATCTCAAGATTCCTTCGAGATCTACATGTGGTCAGTCCAAACCGGTAAACTGCTTGACA 1429
QY 497 alLeuThr---TyrThrAsnPro---SerTyrProAsn---LeuAspHisGlnArgSerV 514
Db 1430 TCCTCACTGCCATACCGCTCTCTATATCAGG-CCTCGCTTCCTCCACCGGTAACTCAG 1488
QY 514 al---LeuPheIleAsnLysLysTyrPheLeuValIleAspArgAlaIleGlyGluAlar 533
Db 1489 TTGGCATCTCTCTGCGATCGTTCATCC-GTT-----TATGGTCAGTCTT 1535
QY 533 hrGlyAsnLeuGlyValHisTrpGlnLeuLysGluAspSerAsnProValPheAspLysT 553
Db 1536 CGGCGCAT-----CAAGAGCCACCAGACCGGATGAGCTTCGCGG 1574
QY 553 hrLysAsnArgValTyrThrTyrArgAspGlyAsnAsnLeuMetIleGlnSerLeuA 573
Db 1575 CGAAGCGACTGCTGGCGTTCAGG-CCTGACCGAATGAGATCGCGCTTCTACTTTGA 1633
QY 573 snAlaAspArgThrSerLeuAsnGluGluGluLysValSerTyrValTyr-----A 591
Db 1634 ACGGGAATTTGATCTTTATCGATGTGGAAGAAGACAGATTAAGTCTGTGTTATTGAAGCC 1693
QY 591 snLysGlnLeuLysArgProAlaPheValPheGluLys-----ProLysLysAsnAlaG 609
Db 1694 GAAGAGATATTTCTGGAGGAGAAAGTGTGATCAGCGCTTACAGCTGCCAATAACCGCG 1753
QY 609 lyThrGlnAsnPheValSerIleValTyrProTyrAsp-----G 622
Db 1754 CAAGCAAGTATTTCAACAGTGTATCTACACTGCCGACGGTGTGTTGTGCTGGTGTG 1813
QY 622 lyGlnLysAlaProGluIleSerIleArgGluAsnLysGlyAsnAspPheGluLysGlyL 642
Db 1814 GAAGCAGCAAGTATGTTGTGTGTGATCGGACGGAAGCGGTGATGGTGAAGAAGTTC 1873
QY 642 ysLeuAsnLeuThrLeuThrIleAsnGlyLysGlnGlnLeuVal 656
Db 1874 AGATCAGCAAAATCTCAGCTTGAGCGGTACGCAAGAGATGTTG 1917

RESULT 11
US-10-369-493-46899
; Sequence 46899, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
```

```
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 46899
; LENGTH: 2034
; TYPE: DNA
; ORGANISM: Bacillus subtilis
; US-10-369-493-46899

Alignment Scores:
Pred. No.: 0.00559 Length: 2034
Score: 120.00 Matches: 126
Percent Similarity: 32.77% Conservative: 87
Best Local Similarity: 19.38% Mismatches: 214
Query Match: 3.43% Indels: 224
DB: 16 Gaps: 32

US-09-802-285A-2 (1-659) x US-10-369-493-46899 (1-2034)
QY 41 TyrSerGlyLeuGluLysValAsnLysAlaValAlaAlaGlyAsnTyrAspAspAlaAla 60
Db 313 TTTTCTGGAAGCGCAGTTGTAGATAAAATAACACAGCGGTTTCAACAGGCAAGAG 372
QY 61 LysAlaLeuLeuAlaTyrTyr-----ArgGluLysSerLysAlaArgGluProAsp 77
Db 373 AACCGCTTGTGCGCATTTATACAGGATCGGGAAGGCCATCAAGTGCAGAAAGTATTGCC 432
QY 78 PheSerAsnAlaGluLysProAlaAspIleArgGlnProIleAspLysValThrArgGlu 97
Db 433 TATAGCAACGACAAA-----GGAAGAACATGACGACGAG 465
QY 98 MetAlaAspLysAlaLeuValHisGlnPheGlnProHisLysGlyTyrGlyTyrPheAsp 117
Db 466 TAGCGTGGCAACCTCTGCTATT-----CCAAACCCGGGT 498
QY 118 TyrGlyLysAspIleAsnTrpGlnMetTrpProValLysAspAsnGluValArgTrpGln 137
Db 499 ---AAAAAGATTTT-----CGCATCCAAAGTCTTTGGTAT 534
QY 138 LeuHisArgValLysTrpTrpGlnAlaMetAla-----LeuValTyrHis 152
Db 535 GAGAAAGAAAAAAGTGGGTGATGGTCTTGGCGCGGTGACCGAATCCTCATTTAT 591
QY 153 AlaThrGlyAspGluLysTyrAlaArgGluTrpValTyrGlnTyrSerAspTrpAlaArg 172
Db 592 -----ACATCAAAAAATCTGAAGCAGTGACGCTAT 621
QY 173 LysAsnProLeuGlyLeuSerGlnAspAsnAspLysPheValTrpArgProLeuGluVal 192
Db 622 GCAAGTCAATTTGGACAGGATCAAGNAGCCACGGGGAGTATGG----- 666
QY 193 SerAspArgValGlnSerLeuProProThrPheSerLeuPheValAsnSerProAlaPhe 212
Db 667 -----GAAATGCCGGATTTATTGAGCTTCCGCTAGACGCAATCCGAAT 711
QY 213 ThrProAlaPheLeuMetGluPhe-----LeuAsn 222
Db 712 CAAAGAAATGGGTCTATGTCAGGTGTCGGAACGAGCGGTCTCGGAGGATCAGGC 771
QY 223 SerTyrHisGlnGlnAlaAspTyrLeuSerThrHisTyrAlaGluGlnGly-----Asn 240
Db 772 ATGCAATATTTTGTAGGGGACTTTGATGGAACCTCAITTTAAAAATGAAACCCGCCAAC 831
QY 241 HisArgLeuPhe---GluAlaGlnArgAsnLeuPheAlaGlyValSerPheProGluPhe 259
Db 832 AAAATTCTCTGGACGGATTTACGGCAGAGACTTTTATCGCGCTGATCTCGTCTGATATT 891
QY 260 -----LysAspSerProArg----- 264
Db 892 CCATCCACAGACAGCCCGCGCTATGTTAGGTGGATGAGCAATTGGCAATATGCGAAT 951
QY 265 -----TrpArgGlnThrGlyIleSerValLeuAsnThrGluIleLys 278
```



Db	952	GATGTTCCGACATCCCATGGAGAGTGCA-----ACGTCCATCCCAAGAGTAAAA	1005	Db	1935	TGTGATTGATGCACATGATTCCGTATTTTCAGAGGCCCAATTGGCTTGAATGTGTGG--	1992
Qy	279	LysGlnValTyrAlaAspGlyMetGlnPheGluLeuSerProIleTyrHisValAlaAla	298	Qy	541	nLeuLysGluAspSerAsnProValPhe	550
Db	1006	TTGAAGCGTTTACCGAAGGGTTAGAGTGGTCCAAACACCGTGAAAGAGTGGAAACC	1065	Db	1993	-----GACGGGACTGCTGTCTTT	2010
Qy	299	Ile-----AspIle	301	RESULT 12			
Db	1066	ATTCCGGGAACCTTAAGAAGTGAAGAATCTGCACATATCCCTCCCAAGTCATATGTG	1125	US-10-369-493-26002			
Qy	302	PheLeuLysAlaTyrGlySerAlaLysArgValAsnLeuGlu---LysGluPheProGln	320	; Sequence 26002, Application US/10369493			
Db	1126	TTAGCGGGCAATCTGGAGATCCCTATGAATAAATATGCAATTTAAAGTAAGCCCTGGT	1185	; Publication No. US20030233675A1			
Qy	321	Ser-----TyrValGlnThrValGluAsnMetIleMet-----	331	; GENERAL INFORMATION:			
Db	1186	TCAGCTGCTGATTTGGTTTAAAGTCCGAACAGGTGAAATCAATTTACGAAGTCGCG	1245	; APPLICANT: Cao, Yongwei			
Qy	332	-----AlaLeuIleSerLeuProAspTyrAsnThrProMetPhe	346	; APPLICANT: Hinkle, Gregory J.			
Db	1246	TATGACCGAAGAACCCCAATTTGCTTGCAGGAGCGAGTCAGGCAACGACACCTTT	1305	; APPLICANT: Slater, Steven C.			
Qy	347	GlyAspSerTrpIleThrAspLysAsn-----PheArgMetAlaGlnPheAlaSer	363	; APPLICANT: Goldman, Barry S.			
Db	1306	AATCCGGCGCTTAAACACCGGAAGAAAGAACAGCCCGCTGAACCGGTAAATCGGAAGGTT	1365	; APPLICANT: Chen, Xianfeng			
Qy	364	TrpAlaArgValPheProAlaAsnGlnAlaIleLysTyrPheAlaThrAspGlyLysGln	383	; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF			
Db	1366	AAGTTGGCGATTTTGTTCAGCCGCTCTCGGTTGAAGTATTTGGGAATGACGGAAGCAG	1425	; FILE REFERENCE: 38-10(52052)B			
Qy	384	GlyLysAla-----ProAsnPheLeuSerLysAlaLeu-----Ser	395	; CURRENT APPLICATION NUMBER: US/10/369,493			
Db	1426	GTCATAACGGATATTATTCTCCGACGACGATCAAGCAAAAGGCTTCAATTTATGCTGCA	1485	; CURRENT FILING DATE: 2003-02-28			
Qy	396	AsnAlaGly-----	398	; PRIOR APPLICATION NUMBER: US 60/360,039			
Db	1486	AATGGCGGTGAAGGTAAAGTAAATCTTTAAGATACACCCCTTTAAAGAGTATGGGGAAGC	1545	; PRIOR FILING DATE: 2002-02-21			
Qy	399	-----PheTyrThrPheArgSerGlyTrpAspLysAsnAlaThrValMetValLeuLys	416	; SEQ ID NO 26002			
Db	1546	ACACCTTTTATCTCCAATATGACTGGCTGG-----ACGACTGTA-----	1584	; TYPE: DNA			
Qy	417	AlaSerProProGlyGluPheHisAlaGlnProAspAsnGlyThrPheGluLeuPheIle	436	; ORGANISM: Schizosaccharomyces pombe			
Db	1585	-----AATGGCACGTCGTCAGACACACAAT	1608	US-10-369-493-26002			
Qy	437	LysGlyArgAsnPheThrProAspAlaGlyValPheValTyr-----SerGlyAsp	453	Alignment Scores:			
Db	1609	GAGGGAACAAAGGAGTGGACGGCGATTCTTTATCTTCTCTTCAGCATCCGGGTCA	1668	Pred. No.:	0.00649	Length:	1737
Qy	454	Glu-AlaIleMethLysLeuArgAsnTrpTyrArgGln-----	465	Score:	118.50	Matches:	111
Db	1669	GACTTCACTATGATCTGATATCACCATTAAAGATGGAACCGAAGAGGGGCGAGGCA	1728	Percent Similarity:	34.96%	Conservative:	75
Qy	466	-----ThrArgIleHisSerThrLeuThrLeuAspAsnGlnAsnMetVa	480	Best Local Similarity:	20.86%	Mismatches:	195
Db	1729	CTAATGTTTCGCTCGACAAAGATCCCAAAACGGTTACCTTGGCC-AAATGTGGATCGGAA	1787	Query Match:	3.39%	Indels:	151
Qy	480	IleThrLysAlaArgGlnAsnLysTrpGluThrGlyAsnAsnLeuAspValLeuThrTy	500	DB:	16	Gaps:	31
Db	1788	GCATGACCTAGTGAATCTTTTAATTTGAGAACGGTGTCTTCTGTCTATCTGCTGAATA	1847	US-09-802-285A-2 (1-659) x US-10-369-493-26002 (1-1737)			
Qy	500	rThrAsnProSerTyrProAsnLeuAspHisGlnArgSerValLeuPheIleAsnLysLy	520	Qy	112	GlyTyrGlyTyrPheAspTyrGlyLysAspIleAsnTrpGlnMetTrpProValLysAsp	131
Db	1848	CAAAACACCG-----ATAGACGCTTAATAAAAAA	1874	Db	208	GGTTATGATGTTCTTGATTAC---AAACAATGACTCTCGATATGGAACCTTAGAGAT	264
Qy	520	stYrPheLeu-----ValIleAspArgAl	528	Qy	132	---AsnGluValArgTrpGlnLeuHisArgValLysTrpTrpGlnAlaMetAlaLeuVal	150
Db	1875	GATCATCTGAAACACAGAGCGCGGATCGCTTTAAATCTAATTTAGATGATCGTCT	1934	Db	265	TTGGACAGCTTAATGAAAGCTCTACAGAAAGGACATGAGTTAGTTATGACTTGGTG	324
Qy	528	alleGlyGluAla-----ThrGlyAsnLeuGlyValHis---TrpGlu	541	Qy	151	TyrHisAlaThrGlyAspGluLysTyrAlaArgGluTrpValTyrGlnTyrSerAspTrp	170

```
Db 583 CACGATATCTACGTTCTGCTGTAGAGGAGTA-----GATGG-----TTTCGA 630
Qy 290 LeuSerProIleTyrHisValAlaAlaIleAspIlePheLeuLysAla----- 305
Db 631 CTTGATGCTATTAAACATGATCAGCAAGACCAAGGTTTCTAGATGCTCCAAATTTACTGAT 690
Qy 306 -----TyrGlySerAlaLysArgValAsnLeu 314
Db 691 GACAGGTACGAATATCAACTAGCTATCAATATPACGCCAATGTCCTCCAGAAATTCAC-- 747
Qy 315 GluLysGluPheProGlnSerTyrValGlnThrValGluAsnMetIleMetAlaLeuIle 334
Db 748 -----GAGTATCTTAATGGGATAGGCATATT----- 774
Qy 335 SerIleSerLeuProAspTyrAsnThrProMetPheGlyAsp---SerTrpIleThrAsp 353
Db 775 -----CTTACAGAGTATGACGCTTTTCTGTAGGGAAATGCTTACGTTTTCGAT 825
Qy 354 LysAsn-----PheArgMetAlaGln 360
Db 826 ACGAACGAATCTGCATGCTGTGGTGTAGCTAGAGCACTGACGATGATATTCAG 885
Qy 361 PheAlaSerTrpAlaArgValPheProAlaAsnGlnAlaIleLysTyrPheAlaThrAsp 380
Db 886 TTTGATTTTGTGACCTTGATCTCGATCTTAATCAGCAT---AAGTATATTGAAGGAGT 942
Qy 381 GlyLysGlnGlyLysAlaProAsnPheLeuSerLys-----AlaLeuSerAsn 396
Db 943 TGGGAGCTTCTCTGATTTGAAAGATCATTAATAAATGCGCAGTCTGCACTACTTTTCGGT 1002
Qy 397 AlaGlyPhe---TyrThrPheArgSerGlyTyrAspLysAsnAlaThrVal---MetVal 414
Db 1003 GTGGTGTGAATGCTTCCTTATTGAAATCACCACCAACGAGAACTGTTTCACGATAT 1062
Qy 415 LeuLysAlaSerProProGlyGluPheHisAlaGlnProAspAsnGlyThrPheGluLeu 434
Db 1063 CTTTCTGATTTCCCAAGTAGACATATAGCTCCAAATTTGATGGCTCTTTTATCATA 1122
Qy 435 PheIleLysGlyArgAsnPheThrProAspAlaGlyValPheValTyrSerGlyAspGlu 454
Db 1123 TTCCAAGTGGT-----ACCCCA-----TTGTATTCCAGGTCAGGAA 1161
Qy 455 AlaIleMetLysLeuArgAsnTyrTyrArgGlnThrArgIleHisSerThrLeuThrLeu 474
Db 1162 -----CTTGCTCTTGCCAAATATCCACGAGATGGCCGATGATGAATACCTTGATGTT 1215
Qy 475 AspAsnGlnAsnMetValIleThrLysAlaArgGlnAsnLysTrpGluThrGlyAsnAsn 494
Db 1216 GAAACGAGAACTTC-----TGGA--- 1236
Qy 495 LeuAspValLeuThrTyrThrAsnProSerTyrProAsnLeuAspHisGlnArgSerVal 514
Db 1237 -----CTTTTATGATGGCAATCTCTAGTCAGGAAGAAAT-----GAGAAGACTATG 1284
Qy 515 LeuPheIleAsnLysLysTyrPheLeuValIleAspArgAlaIleGlyGluAlaThrGly 534
Db 1285 GACATTGTTACAAACGT-----GCTCGTGACACGGT 1317
Qy 535 AsnLeuGlyValHisTrpGlnLeuLysGluAspSerAsnProValPheAspLysThrLys 554
Db 1318 CGCACACCAATGATTTGG-----GATAGCTCACCCACGGTGTTCACAAAG 1365
Qy 555 -----AsnArgValTyrThrThrTyrArgAspGlyAsnAsnLeuMet 568
Db 1366 GTGGAGTAAACCGGTGATGAGTGACTTAATGACTACAAAGATGGAAT----- 1416
Qy 569 IleGlnSerLeuAsnAlaAspArgThrSerLeuAsnGluGluGlyLysValSerTyr 588
Db 1417 -----GCTGCCAATCAAGTTAATGACCTGAAACCTGAAAGTCTTACACATT 1458
Qy 589 -----ValTyrAsnLysGluLeuLys 595
```

```
Db 1459 TGGTCAAGGCATTGGAGCTTCGTTAAAGAATTGAAG 1494
RESULT 13
US-10-369-493-26008
; Sequence 26008, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hankie, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 26008
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-26008
Alignment Scores:
Pred. No.: 0.00649 Length: 1737
Score: 118.50 Matches: 111
Percent Similarity: 34.96% Conservative: 75
Best Local Similarity: 20.86% Mismatches: 195
Query Match: 3.39% Indels: 151
DB: 16 Gaps: 31
US-09-802-285A-2 (1-659) x US-10-369-493-26008 (1-1737)
Qy 112 GlyTyrGlyTyrPheAspTyrGlyLysAspIleAsnTrpGlnMetTrpProValLysAsp 131
Db 208 GGTATGATGTTTCTGATTTAC---AAACAATTGACTCTCGATATGGAATTTAGAGAT 264
Qy 132 ---AsnGluValArgTrpGlnLeuHisArgValLysTrpTrpGlnAlaMetAlaLeuVal 150
Db 265 TTGCACAGGCTAATAAGAGCTCTACACGAAAGGACATGAAGTTAGTTATGGACTTGGTG 324
Qy 151 TyrHisAlaThrGlyAspGluLysTyrAlaArgGluTrpValTyrGlnTyrSerAspTrp 170
Db 325 TTAACCCACACTTCTGATCAA-----CATGAGTGG---TTTAGAGGAGTCCAGATCT 372
Qy 171 AlaArgLysAsnProLeuGlyLeuSerGlnAspAsnAspLysPheValTrpArgProLeu 190
Db 373 TCTAAACAATCCG-----AAGCGAGATTGGTACTTTTGGAAACCACT 417
Qy 191 GluValSerAspArgValGlnSerLeuProProThrPheSerLeuPheValAsnSerPro 210
Db 418 AGGTACAATGAAAGGCGAAGCGTTTACCCCA----- 450
Qy 211 AlaPheThrProAlaPheLeuMetGluPheLeuAsnSerTyrHisGlnGlnAlaAspTyr 230
Db 451 -----ATAATTTGGAGAACTATTTCGATACT 477
Qy 231 LeuSerThrHisTyrAlaGluGlnGlyAsnHisArgLeuPheGluAlaGlnArgAsnLeu 250
Db 478 TCGCGCTGGGAATGGGACGAAGCTACGCAA-----GAATACTACTACATCTT 525
Qy 251 PheAlaGlyValSerPheProGluPheLys---AspSerProArgTrpArgGlnThrGly 269
Db 526 TGTGCC---GTAGGGCAACCCGACTTAAATTGGGAAACCCCAAGTTAGGAAACCGGTA 582
Qy 270 IleSerValLeuAsnThrGluIleLysLysGlnValTyrAlaAspGlyMetGlnPheGlu 289
Db 583 CACGATATCTAGTTCTGCTTGTAGAGGAGTA-----GATGG-----TTTCGA 630
Qy 290 LeuSerProIleTyrHisValAlaAlaIleAspIlePheLeuLysAla----- 305
```



5278 TATCAAAAAATTTACGAAGAGTATGGAGATATGGCGACTTTTTCATCAATTTACTTCAAGGT 5337  
QY  
138 LeuHisArgValLysTrpTrpGlnAlaMetAlaLeuValTyrHisAlaThrGlyAspGlu 157  
QY  
5338 TTACACGACCGCGAAATGAAACTTGTAAATGGATCTTGTCTTAAACCATACTTCTGATGAA 5397  
Db  
158 LysTyrAlaArgGluTrpValTyrGlnTyrSerAspTrpAlaArgLysAsnProLeuGly 177  
QY  
5398 -----CACAGTGG-----TTGAGGAAATCGAAAAATCCAAAGCAACCGGTAC--- 5442  
Db  
178 LeuSerGlnAspAsnLysPheValTyrArgProLeuGluValSerAspArgValGln 197  
QY  
5443 -----AGAGATTACTATTCTCGCGCGAG-----GAAAAACGAAATCAAT 5481  
Db  
198 SerLeuProProThrPheSerLeuPheValAsnSerProAlaPheThrProAlaPheLeu 217  
QY  
5482 AACTGGGGCTCGATTCTTCAG-----GGTCCAGCATGGGAA-----TTG 5520  
Db  
218 MetGluPheLeuAsnSerTyrHisGlnGlnAlaAspTyrLeuSerThrHisTyrAlaGlu 237  
QY  
5521 GACGAAAAACAGGTGAGTAC-----TATTTACACCTGTTCTCCAAAAA 5565  
Db  
238 GlnGly-----AsnHisArgLeuPheGluAlaGlnArgAsnLeuPheAlaGlyValSer 255  
QY  
5566 CAACCTGATTAAACTGGGAAACCCCTAAATTCGCCAAGATGATATATATATGATGAAA 5625  
Db  
256 PheProGluPheLysAspSerProArgTrpArgGlnThrGlyLe----- 270  
QY  
5626 TTC-----TGCTAGATAAAGCATTTGATGGCTTCCCGCATG 5661  
Db  
271 SerValLeuAsnThrGluLeLysLysGlnValTyrAlaAspGlyMetGlnPheGluLeu 290  
QY  
5662 GAGGTAAATTTATTTCCAAAAATCTGATTTCCAGATGTCCTCAGTACAGAT--- 5718  
Db  
291 SerProIleTyrHisValAlaAlaIleAspIlePheLeuLysAlaTyrGlySerAlaLys 310  
QY  
5719 GGTCAAAATTTATGGCGATGCTGGCAATGAT-----TTTTGTAATGGACCG 5763  
Db  
311 ArgValAsnLeuGluLysGluPheProGlnSerTyrValGlnThrValGluAsnMetIle 330  
QY  
5764 CGGATTCAT-----GAGTCTTCGCAAGAAATGAACCAAGATTAATCTCTTAAATAC 5814  
Db  
331 MetAlaLeuIleSerIleSerLeuProAspTyrAsnThrProMetPheGlyAspSerTrp 350  
QY  
5815 GATGTGATGACTGTGGCGAAATGCCCTGGCTAGTACA----- 5853  
Db  
351 IleThrAspLysAsnPheArgMetAlaGlnPheAlaSerTrpAlaArgValPhe----- 368  
QY  
5854 ---ACCGAC-----GGCCAAATTTACACGAAT 5877  
Db  
369 ProAlaAsnGlnAlaIleLysTyrPheAlaThr----- 379  
QY  
5878 CCAGCAATATGAAGTCGATATGATTTTACATTTGAACATATGAACCTAGATTCGAT 5937  
Db  
380 ---AspGlyLysGlnGlyLysAlaProAsnPheLeuSerLysAlaLeuSerAsnAlaGly 398  
QY  
5938 GCGACAAATAATGGGATCTAAAACCAATCTACTTACCAGATTTAAAAGAAAATATGTCT 5997  
Db  
399 PheTyrThrPhe-----ArgSerGlyTrpAspLysAsnAlaThrValMetValLeu 415  
QY  
5998 GAATGGCAAGTCGGCTTCAAGAAAACGGTTGGAAAT----- 6033  
Db  
416 LysAlaSerProProGlyGluPheHisAlaGlnPro----- 427  
QY  
6034 -----AGCTTATATGGAATATCATGACCAACCGGTATCGTTTCTCGTTTGGGAAT 6087  
Db  
428 AspAsn-----GlyThrPheGluLeuPheIleLys 437  
QY  
6088 GACAAATCGTTTCCGTGTTCTGTCAGCAAAATGCTTGCACACTGTTTGCATATGATGAAA 6147  
Db  
438 GlyArgAsnPheThrProAspAlaGlyValPheValTyrSerGlyAspGluAlaIleMet 457  
QY  
6148 GGA-----ACGCCA-----TATATTTATCAAGTGAAGAAATTTGGGATG 6186  
Db

458 LysLeuArgAsnTrpTyrArgGlnThrArgIleHisSerThrLeuThrLeuAspAsnGln 477  
QY  
6187 -----ACAAATGTTCAITTTTGA-----ACGCTAGATGATTAC 6219  
Db  
478 -----AsnMet----- 479  
QY  
6220 CGTGATATCGAAACGCTAAACATGTACAAAGAGCGTAAGGAGCAGACACAGCATGAA 6279  
Db  
480 -----ValIleThrLysAlaArgGlnAsn-----LysTrp 489  
QY  
6280 AGCATCATGTCAGTCGATTATACAAAAGGCGGTGACAAATCGAGACGCGTACCACTGG 6339  
Db  
490 GluThrGlyAsnAsnLeuAspValLeuThrTyrThr-----AsnProSer 504  
QY  
6340 GATAATAGCGAGATGCTGGCTTCACAAACCGAACCGCTTGGCTTAAAGTCAACCCCGCG 6399  
Db  
505 TyrProAsnLeuAspHisGlnArgSerValLeuPheIleAsnLysLysTyrPhe----- 522  
QY  
6400 TATACAGAAATCAACAAACGAAGAGCGCTGAAAAACCCAGATTCATTTCTACTACTAC 6459  
Db  
523 -----LeuValIleAspArgAlaIleGlyGlu-----AlaThrGlyAsnLeuGlyVal 538  
QY  
6460 CAAATCTTATTAAGTTTACGAAAAACGACTGAAATCATCAACTGGTAATTTATCGTTTG 6519  
Db  
539 HisTrpGlnLysGluAspSerAsnProValPheAspLysThrLysAsnArgValTyr 558  
QY  
6520 -----TTATTAACCTAAAGATGAAGCGATTTTC-----GCATAT 6552  
Db  
559 ThrThrTyrArgAspGlyAsnAsnLeuMetIleGlnSerLeuAsnAlaAspArgThrSer 578  
QY  
6553 GAACGCTACACAGAAAATGAAAAATTAGTGGTT-----TTATGTAAT 6594  
Db  
579 LeuAsnGluGluGluGlyLysValSer-----TyrValTyrAsnLysGluLeuLysArg 596  
QY  
6595 TTCACAGAAGAAACAAAGTTATTTCTGATGAAAAACATTTTGACAGAAATCCAAAAAGGT 6654  
Db  
597 ProAlaPheValPheGluLysProLysLysAsnAlaGlyThr---GlnAsnPheValSer 615  
QY  
6655 TCCGTGCTGTAACAATGTTCTTAATATATAGAGGAACCTTTACGACCTTACGAAGCA 6714  
Db  
616 IleValTyr 618  
QY  
6715 ATCGTTTAT 6723  
Db

RESULT 15  
US-10-335-977-4204  
; Sequence 4204, Application US/10335977  
; Publication No. US20040052799A1  
; GENERAL INFORMATION:  
; APPLICANT: DOUGLAS SMITH et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
; RELATING TO HELICOBACTER PYLORI FOR  
; DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 10031  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: LAHIVE & COCKFIELD  
; STREET: 28 State Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02109-1875  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: Windows NT 4.0  
; SOFTWARE: UNIX  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/335,977  
; FILING DATE: 30-Dec-2002  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/993,002  
; FILING DATE: 17-DEC-1997

ATTORNEY/AGENT INFORMATION:  
 NAME: Mandragoras, Amy E.  
 REGISTRATION NUMBER: 36,207  
 REFERENCE/DOCKET NUMBER: GTN-018  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 227-7400  
 TELEFAX: (617) 742-4214  
 INFORMATION FOR SEQ ID NO: 4204:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 4455 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: circular  
 MOLECULE TYPE: DNA (genomic)  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Helicobacter pylori  
 FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (B) LOCATION 1...4455  
 SEQUENCE DESCRIPTION: SEQ ID NO: 4204:  
 US-10-335-977-4204

Alignment Scores:  
 Pred. No.: 0.0379 Length: 4455  
 Score: 117.50 Matches: 123  
 Percent Similarity: 33.53% Conservative: 100  
 Best Local Similarity: 18.50% Mismatches: 254  
 Query Match: 3.36% Indels: 189  
 DB: 13 Gaps: 29

US-09-802-285A-2 (1-659) x US-10-335-977-4204 (1-4455)

QY 86 AspileArgInProIleAspLysValThrArgGluMetAla-----AspLysAlaLeu 103  
 DB 1798 GATATTGCAAGCGGACGCGCTTTAAACACGATGTAGCGCTTTGGCGAGCAATG 1857  
 QY 104 ValHisGlnPheGlnProHisLysGlyTyrGlyTyrPheAspTyrGlyLysAsp---Ile 122  
 DB 1858 ATTGCGGAATTT-----TTAGGCCAAGACACGCTC 1887  
 QY 123 AsnTrpGlnMetTrpProValLysAspGlnValArgTrpGlnLeuHisArgValLys 142  
 DB 1888 AATCTTTAGAAAGCTGTGCAAAACGACGAGATTAAGAGCGTTTATAGACAAAGTCTTA 1947  
 QY 143 TrpTrpGlnAlaMetAlaLeuValTyr---HisAlaThrGlyAsp-----GluLysTyr 159  
 DB 1948 GCGGCTAAAGGATTAGGCTTATTATGACAAAGGTTTGGGGGATTGATCCCTAATCTT 2007  
 QY 160 AlaArgGluTrpValTyrGlnTyrSerAspTrpAlaArgLysAsnProLeuGlyLeuSer 179  
 DB 2008 GGTAAAGGGGATTTTCGCT-----CCCTATGGCTTGAGT 2043  
 QY 180 GlnAspAsnAspLysPheValTrpArgProLeuGluValSerAspArgValGlnSerLeu 199  
 DB 2044 CAA-----GTGGGCAAAAGGGGATTTAGTTTCAACGCGCAGC--- 2085  
 QY 200 ProProThrPheSerLeuPheValAsnSerProAlaPheThrProAla-----PheLeu 217  
 DB 2086 -----AATGTTTGTGCAAAATTCACCTTCTCTAACGCTAATGAGGACG 2133  
 QY 218 MetGluPhe----- 220  
 DB 2134 CTCAGTTTAAACGCGAGAAATTCGCTCAATTTTTCGCGGAAACAACCATCGCTTTTCACT 2193  
 QY 221 -----LeuAsnSerTyrHisGlnGlnAlaAspTyrLeuSerThrHisTyr 235  
 DB 2194 AACCAATTCGGAACGCTCAATTTGTGCTAATCAAGTTTCTAACTAATCAACGTCACCATG 2253  
 QY 236 -----AlaGluGlnGly-----AsnHisArgLeuPheGluAlaGln 247  
 DB 2254 CTTAACGCTAGCAACGCGCTTAAGATTAACGCCACTAATAACAAATGTTTCCGTGCTCA 2313

QY 248 ArgAsnLeuPheAlaGlyValSerPheProGluPheLysAspSerProArgTrpArgGln 267  
 DB 2314 GGCAATCTGTTTATCAACGCTAGCTCGGTGCAACAAAGCGATCCAACGACAGCTAGCGCC 2373  
 QY 268 ThrGly----- 269  
 DB 2374 ACAAAACCCCTTGACACCCGCTCAAAATAACGCTTCTTAGTAATGCGTCAACACACGCG 2433  
 QY 270 ---IleSerValLeuAsnThrGluLysLysGlnValTyrAlaAspGlyMetGlnPhe 288  
 DB 2434 CCAATCGCTTAAATAATAACGATCAAAAGCTTGGTGTACGCGAATGCTTCAATTT 2493  
 QY 289 GluLeuSerProIleTyrHisValAlaIleAspIlePheLeuLysAlaTyrGlySer 308  
 DB 2494 TCA---GGCAATATTTACGCTAACCGGGTGTGTAT---TTTCAAAATTAAGGCTCT 2547  
 QY 309 Ala----- 309  
 DB 2548 GCAAAACGTTAAACACCTGTATCTTTACAATAACGCTCAATTCCAAGCCAAACACCTCAG 2607  
 QY 310 -----LysArgValAsnLeu 314  
 DB 2608 ATTTCCAACCAACGCGTATTAGAGAAACGCTAGCTTTGTAAACGAATAACATTAAACATT 2667  
 QY 315 GluLysGluPheProGlnSerTyrValGlnThrValGlu-----AsnMetIleMet 331  
 DB 2668 CAAGGAGCGCTTTAACAACAACGCCACGCAAAATAAGAGGTCTTCAAAATTTAGTGATC 2727  
 QY 332 AlaLeu---IleSerIleSerLeuProAspTyrAsnThrProMetPheGlyAspSerTrp 350  
 DB 2728 GCTTCAAAACGCTTCTTTAAGCACCGGATTTATGGGTAGAAAGTAGGGGGG----- 2778  
 QY 351 IleThrAspLysAsnPheArgMetAlaGlnPheAlaSerTrpAlaArgValPheProAla 370  
 DB 2779 ---GCATTCAATAATTTGGAGCGATCCATTTTAAATTAGAAATTTCTCAACGCTGTGA 2835  
 QY 371 AsnGlnAlaIleLysTyrPheAlaThrAspGlyLysGlnGlyLysAlaProAsnPheLeu 390  
 DB 2836 AATCGCTCAATCAAGTAGGGGGATCATTAATCTCAACACCCCAACGCTTTTATG 2895  
 QY 391 SerLysAlaLeuSerAsnAlaGlyPheTyrThrPheArgSerGlyTrpAspLysAla 410  
 DB 2896 AATGTCAGCGTGGCTAATGGCGAACTTACACT----- 2928  
 QY 411 ThrValMetValLeuLysAlaSerProGlyGluPheHisAlaGlnProAspAsn--- 429  
 DB 2929 -----TTATTAAGACGCGCTTATATTGATTACATATCAACCTTAACAGCTTG 2979  
 QY 430 ---GlyThrPheGluLeuPhe-----IleLysGlyArgAsnPheThrProAsp 444  
 DB 2980 CAATCGTATTTGAAGCTCTATACCTTAATCAATATCAACGGAACACACATAGAGAAAA 3039  
 QY 445 AlaGlyValPheValTyrSerGlyAspGluAlaIleMet---LysLeuArgAsnTyrTrp 464  
 DB 3040 AACGCGCTTATGACTTATTTGGGCCAACGGGTTTATTATACAGATAAGGGTTATTATG 3099  
 QY 464 GlnThrArgIleHisSerThrLeuThrLeuAspAsnGlnAsnMetValIleThrLysAl 484  
 DB 3100 -AGTGTAGCACTACCTAATCAACACGCTCTCAAAACAACATTTTAAAGCCTTCTGT 3158  
 QY 484 aArgGlnAsnLysTrpGluThrGlyAsnAsnLeuAspValLeuThrTyrThrAsnProSe 504  
 DB 3159 CTTCAACACAGATTAATAATGTCTTATGGTAATAAAGTGATGGATTTACCCCTCCAC 3218  
 QY 504 rTyrProAsnLeuAspHisGlnArgSerValLeuPheIleAsnLysLysTyrPheLeuVa 524  
 DB 3219 CTTA-----CAGGATTACATTGTGGGCAATTCAAGACAAAGCGCACTCAA 3263  
 QY 524 lIleAspArgAlaIleGlyGluAlaThrGlyAsnLeuGlyValHisTrp----- 540  
 DB 3264 TCMAATTCAAGCTGTGG-----GGGAATAACGCTATCAAGTGGCTTTCACATT 3314

```
QY 541 -GlnLeuLysGluAspSerAsnProValPheAspLysThrLysAsnArgValTyrThrTh 560
Db : : : : : ||| ||| : : : |||
3315 GATGATGGAGACTAAAGAAACCCGCTTTT----- 3345
QY 560 rTyrArgAspGlyAsnAsnLeuMetIleGlnSerLeuAsnAlaAspArgThrSerLeuAs 580
Db : : : : : ||| ||| : : : |||
3346 -----GCGCGGATTATTAGAAAAACCACCTTTAAA 3377
QY 580 nGluGluGluGlyLysValSerTyrValTyrAsnLysGluLeuLysArgProAlaPheVa 600
Db ||| ||| ||| : : : ||| : : : |||
3378 TGAATCTTTAGC-----GTAACAAAAGATCTTCAAAACACCCGCAAGCTT 3422
QY 600 lPheGluLysproLys---LysAsnAlaGlyThrGlnAsnPheValSerIleValTyrPr 619
Db : ||| : : : ||| : : : |||
3423 GATTCTAACCTTAATTTTGAATAATACGCTACCGCCTTTTAGAATGGCGAGTTACAC 3482
QY 619 oTyrAspGlyGlnLysAlaProGluIleSer-----IleArgGluAsnLysGlyAs 636
Db : : : : : ||| ||| : : : |||
3483 CCAACAAACCAGCGGTTTGACAAAACCTCTCTGATTTTAGGGCTAGAGAGGAGAGTCCAA 3542
QY 636 nAspPheGluLys 640
Db | ||| : : :
3543 TTTTTCAGAGCGC 3555
```

Search completed: July 31, 2004, 20:35:30  
Job time : 2362 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: July 31, 2004, 15:20:31 ; Search time 138 seconds  
(without alignments)

2650.092 Million cell updates/sec

Title: US-09-802-285A-2

Perfect score: 3494

Sequence: 1 MTTKFKRIIVFAVIALSSG.....KGKLNLTITNGKQOLVLP 659

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Zgapop 6.0 , Zgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO spool p/US09802285/runat\_27072004\_155104\_9063/app query.fasta\_1.839  
-DB=Issued Patents NA -QFWT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFWT=pro -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09802285 @CGN 1 1 128 @runat\_27072004\_155104\_9063 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq:\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3494	100.0	1980	1 US-08-258-639A-3	Sequence 3, Appli
2	3494	100.0	1980	2 US-08-900-951-3	Sequence 3, Appli
3	3494	100.0	1980	5 PCT-US95-07391A-3	Sequence 3, Appli
c	176	5.0	3763	4 US-08-961-527-186	Sequence 186, App
5	166.5	4.8	1950	4 US-09-134-000C-1897	Sequence 1897, Ap
6	132.5	3.8	2097	4 US-09-134-001C-795	Sequence 795, App
7	123.5	3.5	3414	1 US-07-973-320-1	Sequence 1, Appli
8	123	3.5	1542	1 US-08-328-962-1	Sequence 1, Appli
c	120.5	3.4	4403765	3 US-09-103-840A-2	Sequence 2, Appli
c	120.5	3.4	4411529	3 US-09-103-840A-1	Sequence 1, Appli
11	117.5	3.4	3087	4 US-09-543-681A-3009	Sequence 3009, Ap
c	116	3.3	4702	4 US-08-956-171E-268	Sequence 268, App

13	116	3.3	13993	4 US-09-220-132-20	Sequence 20, Appli
14	115.5	3.3	3414	1 US-07-973-320-3	Sequence 3, Appli
15	115.5	3.3	3504	1 US-08-620-717A-8	Sequence 8, Appli
16	115	3.3	2859	5 PCT-US96-05320A-637	Sequence 637, App
17	115	3.3	16535	4 US-08-961-527-74	Sequence 74, Appli
c	115	3.3	1830121	4 US-09-557-884-1	Sequence 1, Appli
c	115	3.3	1830121	4 US-09-643-990A-1	Sequence 1, Appli
19	114.5	3.3	3504	1 US-08-485-568A-5	Sequence 5, Appli
20	114.5	3.3	3504	2 US-08-590-554A-5	Sequence 5, Appli
21	114.5	3.3	3504	2 US-09-184-223-5	Sequence 5, Appli
22	114.5	3.3	1830121	4 US-09-557-884-1	Sequence 1, Appli
23	114.5	3.3	1830121	4 US-09-643-990A-1	Sequence 1, Appli
24	114.5	3.2	4765	1 US-08-750-532-8	Sequence 8, Appli
25	113.5	3.2	4765	3 US-08-894-818B-7	Sequence 7, Appli
26	113.5	3.2	4765	4 US-09-445-472-5	Sequence 5, Appli
27	113.5	3.2	2339	1 US-08-258-639A-1	Sequence 1, Appli
28	112.5	3.2	2339	2 US-08-900-951-1	Sequence 1, Appli
29	112.5	3.2	2339	5 PCT-US95-07391A-1	Sequence 1, Appli
30	112.5	3.2	2784	4 US-09-134-001C-1994	Sequence 1994, Ap
31	111	3.2	2784	4 US-09-976-594-1011	Sequence 1011, Ap
32	110	3.1	13930	4 US-09-108-006C-2	Sequence 2, Appli
33	110	3.1	14070	4 US-09-134-000C-2728	Sequence 2728, Ap
c	109.5	3.1	4320	4 US-09-198-452A-1	Sequence 1, Appli
35	109.5	3.1	1230025	4 US-08-816-346-3	Sequence 3, Appli
36	109	3.1	2858	3 US-08-816-346-57	Sequence 57, Appli
37	109	3.1	2858	3 US-09-335-411-3	Sequence 3, Appli
38	109	3.1	2858	3 US-08-335-411-57	Sequence 57, Appli
39	109	3.1	2858	2 US-08-788-674-7	Sequence 7, Appli
40	109	3.1	2859	2 US-09-604-694B-1	Sequence 1, Appli
41	109	3.1	32798	4 US-08-735-609-4	Sequence 4, Appli
42	109	3.1	34303	2 US-08-735-609-4	Sequence 4, Appli
43	109	3.1	34303	2 US-09-315-372-4	Sequence 4, Appli
44	109	3.1	34303	3 US-09-244-752-4	Sequence 4, Appli
45	109	3.1	34303	3 US-09-244-752-4	Sequence 4, Appli

#### ALIGNMENTS

#### RESULT 1

US-08-258-639A-3

; Sequence 3, Application US/08258639A

; Patent No. 5681733

; GENERAL INFORMATION:

; APPLICANT: Su, Hongsheng

; APPLICANT: Blain, Francoise

; APPLICANT: Bennett, Clark

; APPLICANT: Gu, Kangfu

; APPLICANT: Zimmermann, Joseph

; APPLICANT: Musil, Roy

; TITLE OF INVENTION: Nucleic Acid Sequences And Expression

; TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From

; TITLE OF INVENTION: Flavobacterium heparinum

; NUMBER OF SEQUENCES: 26

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hale and Dorr

; STREET: 1455 Pennsylvania Avenue, N.W.

; CITY: Washington, D.C.

; COUNTRY: U.S.A.

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/258,639A

; FILING DATE: 10 JUNE 1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Healey, William J.

; REGISTRATION NUMBER: 36,160

; REFERENCE/DOCKET NUMBER: 104385.116

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)942-8400

; TELEFAX: (202) 942-8484  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1980 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-258-639A-3

## Alignment Scores:

Pred. No.: 0 Length: 1980  
Score: 3494.00 Matches: 659  
Percent Similarity: 100.00% Conservativeness: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 1 Gaps: 0

US-09-802-285A-2 (1-659) x US-08-258-639A-3 (1-1980)

QY 1 MetThrThrLysIlePheLysArgIleIleValPheAlaValIleAlaLeuSerSerGly 20  
DB 1 ATGACTACGAAATTTTAAAGGATCATTTGTTGCTGTAATTCGCCCTATCGTCGGGA 60  
QY 21 AsnIleLeuAlaGlnSerSerIleThrArgLysAspPheAspHisIleAsnLeuGlu 40  
DB 61 AATATACTTGCACAAAGCTCTTCCATTACAGGAAAGATTTTGCCACATCAACCTTGAG 120  
QY 41 TyrSerGlyLeuGluLysValAsnLysAlaValAlaAlaGlyAsnTyrAspAlaAla 60  
DB 121 TATTCCGGACTGGAAAGGTTAATAAGCAGTTGCTCGCGCAACTATGACGATCGGGC 180  
QY 61 LysAlaLeuLeuAlaTyrArgGluLysSerLysAlaArgGluProAspPheSerAsn 80  
DB 181 AAGACATTACTGGCTACTACAGGAAAAAGTAAGCCAGGAACTGATTTTCAGTAAT 240  
QY 81 AlaGluLysProAlaAspIleArgGlnProIleAspLysValThrArgGluMetAlaAsp 100  
DB 241 GCAGAAAGCTCGCGATATACGCCAGCCATAGATTAAGTTACGCGTGAATGGCCGAC 300  
QY 101 LysAlaLeuValHisGlnPheGlnProHisLysGlyTyrGlyTyrPheAspTyrGlyLys 120  
DB 301 AAGGCTTTGTCCACCAGTTTCAACCGCACAAAGGCTACGGCTATTTTGATTATGGTAAA 360  
QY 121 AspIleAsnTrpGlnMetTrpProValLysAspAsnGlnValArgTrpGlnLeuHisArg 140  
DB 361 GACATCACTGGCAGATGTGCCGCTGTAAGACAAATGAAGTACGCTGGCAGTTGCACCGT 420  
QY 141 ValLysTrpTrpGlnAlaMetAlaLeuValTyrHisAlaThrGlyAspGluLysTyrAla 160  
DB 421 GTAAATGTGTGGCAGGCTATGGCCCTGGTTTATCACGCTACGGCGCATGAAATATGCA 480  
QY 161 ArgGluTrpValTyrGlnTyrSerAspTrpAlaArgLysAsnProLeuGlyLeuSerGln 180  
DB 481 AGAATATGGGTATATCAGTACACCGATTTGGCGCAGAAAAACCAATTTGGGCCCTGTGCGCAG 540  
QY 181 AspAsnAspLysPheValTyrArgProLeuGluValSerAspArgValGlnSerLeuPro 200  
DB 541 GATAATGATAAATTTGTGTGGCGGCCCTTGAAGTGTCCGACAGGGTACAAAGTCYTCCC 600  
QY 201 ProThrPheSerLeuPheValAsnSerProAlaPheThrProAlaPheLeuMetGluPhe 220  
DB 601 CCAACCTTCAGCTTATTGTGTAATCTGCCAGCCTTTTACCCGAGCCTTTTAAATGGAATTT 660  
QY 221 LeuAsnSerTyrHisGlnGlnAlaAspTyrLeuSerThrHisTyrAlaGluGlnGlyAsn 240  
DB 661 TTAACAGTTTACCAACAGGCGGATTTATTTATCTACGCATTATGCGCAACAGGGAAC 720  
QY 241 HisArgLeuPheGluAlaGlnArgAsnLeuPheAlaGlyValSerPheProGluPheLys 260  
DB 721 CACCGTTTATTGAAAGCCCAACCACTGTTTTCAGGCGTATCTTTCCCTGAAATTTAAA 780  
QY 261 AspSerProArgTrpArgGlnThrGlyLysSerValLeuAsnThrGluLysLysGln 280

DB 781 GATTCCACAGATGGAGGCAACCGGCATATCGTGTGACACCGAGATCAAAAAACAG 840  
QY 281 ValTyrAlaAspGlyMetGlnPheGluLeuSerProIleTyrHisValAlaAlaIleAsp 300  
DB 841 GTTTATGCCGATGGGATGCAGTTTGAACCTTACCAATTTTACCATGTAGTGGCATCGAT 900  
QY 301 IlePheLeuLysAlaTyrGlySerAlaLysArgValAsnLeuGluLysGluPheProGln 320  
DB 901 ATCTTCTTAAAGGCCCTATGGTTCTGCAAAACGAGTTAACTTAAAAAGAAATTTCCGCAA 960  
QY 321 SerTyrValGlnThrValGluAsnMetIleMetAlaLeuIleSerIleSerLeuProAsp 340  
DB 961 TCTTATGTACAAACTGTAGAAATATAGTATTGGCGCTGATCAGTATTTCACCTGCCAGAT 1020  
QY 341 TyrAsnThrProMetPheGlyAspSerTyrIleThrAspLysAsnPheArgMetAlaGln 360  
DB 1021 TATAACACCCCTATGTTTGGAGATTATGATGATTACAGATAAAAATTTTCAGATGGCACAG 1080  
QY 361 PheAlaSerTrpAlaArgValPheProAlaAsnGlnAlaIleLysTyrPheAlaThrAsp 380  
DB 1081 TTTGCCAGCTGGGCCCGGTTTTCCCGCAAAACAGGCCATAAAATATTITGCTACAGAT 1140  
QY 381 GlyLysGlnGlyLysAlaProAsnPheLeuSerLysAlaLeuSerAsnAlaGlyPheTyr 400  
DB 1141 GGCACAAACAGGTAAAGCGCCCTAACTTTTATCCAAAGCATTTGAGCAATGCAGGCTTTAT 1200  
QY 401 ThrPheArgSerGlyTrpAspLysAsnAlaThrValMetValLeuLysAlaSerProPro 420  
DB 1201 ACOTTTAGAAAGCGGATGGGATAAAAAATGCAACGGTTATGTGTTATTAAGGCAGTCTCTCC 1260  
QY 421 GlyGluPheHisAlaGlnProAspAsnGlyThrPheGluLeuPheIleLysGlyArgAsn 440  
DB 1261 GGGGAATTTTATGCCAGCCGGAATAACGGGACITTTTGAACITTTTATAAGGGCAGAAAC 1320  
QY 441 PheThrProAspAlaGlyValPheValTyrSerGlyAspGluAlaIleMetLysLeuArg 460  
DB 1321 TTTTACCCAGACCGCGGGTATTGTGTATAGCGGCGACGAAGCCATCATGAAATCGCG 1380  
QY 461 AsnTrpTyrArgGlnThrArgIleHisSerThrLeuThrLeuAspAsnGlnAsnMetVal 480  
DB 1381 AACTGTACTCGTCAAAACCCGCTATACAGCAGCCTTACACTCGCAATCAAAATATGGTC 1440  
QY 481 IleThrLysAlaArgGlnAsnLysTyrGluThrGlyAsnAsnLeuAspValLeuThrTyr 500  
DB 1441 ATTTACAAAGCCCGGCAAAACAAATGGGAAACAGGAAATAACCTTGATGTGCTTACCTAT 1500  
QY 501 ThrAsnProSerTyrProAsnLeuAspHisGlnArgSerValLeuPheIleAsnLysLys 520  
DB 1501 ACCAACCCAGCTATCCGAATCTGGACCATCGCGCAGTGTACTTTTTCATCAACAAAAA 1560  
QY 521 TyrPheLeuValIleAspArgAlaIleGlyGluAlaThrGlyAsnLeuGlyValHisTrp 540  
DB 1561 TACTTTCTGCTCATCATAGGCAATAGGCAAGTACCGGAAACCTCGGCGCTACACTGG 1620  
QY 541 GlnLeuLysGluAspSerAsnProValPheAspLysThrLysAsnArgValTyrThrThr 560  
DB 1621 CACCTTAAAGAACACAGCAACCTCTGTTTCGATAAGCAAAAGAACCGGGTTTACACCACT 1680  
QY 561 TyrArgAspGlyAsnAsnLeuMetIleGlnSerLeuAsnAlaAspArgThrSerLeuAsn 580  
DB 1681 TACAGAGATGGTAACCACTGATGATCCCAATCGTTGAATGCGGACAGGACCGCTCAAT 1740  
QY 581 GluGluGluGlyLysValSerTyrValTyrAsnLysGluLeuLysArgProAlaPheVal 600  
DB 1741 GAAGAAGAGGAAGATATCTTATGTTTCAATAGAGCTGAAAGACCTGCTTTCGTA 1800  
QY 601 PheGluLysProLysLysAsnAlaGlyThrGlnAsnPheValSerIleValTyrProTyr 620  
DB 1801 TTTGAAAGGCTAAAAAGAAATGCGGCACACAAAAATTTTGTAGTAGTATTATCCATAC 1860  
QY 621 AspGlyGlnLysAlaProGluIleSerIleArgGluAsnLysGlyAsnAspPheGluLys 640



Db 1861 GACGGCCAGAGGCTCCAGAGATCAGATACGAGGAAAACAGGGCAATGATTTTGAGAAA 1920  
QY 641 GlyLysLeuAsnLeuThrLeuThrLeuThrLeuAsnGlyLysGlnGlnLeuValLeuValPro 659  
Db 1921 GGCAGCTTAATCTTAACCTTACCATTACGAGGAAAACACAGCTTGTTGGTTCT 1977

RESULT 2  
US-08-900-951-3  
; Sequence 3, Application US/08900951  
; Patent No. 5919693  
; GENERAL INFORMATION:  
; APPLICANT: Su, Hongsheng  
; APPLICANT: Blain, Francoise  
; APPLICANT: Bennett, Clark  
; APPLICANT: Gu, Kangfu  
; APPLICANT: Zimmermann, Joseph  
; APPLICANT: Musil, Roy  
; TITLE OF INVENTION: Nucleic Acid Sequences And Expression  
; TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From  
; TITLE OF INVENTION: Flavobacterium heparinum  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Hale and Dorr  
; STREET: 1455 Pennsylvania Avenue, N.W.  
; CITY: Washington, D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC DOS/MS DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/900,951  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/258,639  
; FILING DATE: 10 JUNE 1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Healey, William J.  
; REGISTRATION NUMBER: 36,160  
; REFERENCE/DOCKET NUMBER: 104385.116  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)942 8400  
; TELEFAX: (202)942 8484  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1980 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-900-951-3

Alignment Scores:  
Pred. No.: 0 Length: 1980  
Score: 3494.00 Matches: 659  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 2 Gaps: 0

US-09-802-285A-2 (1-659) x US-08-900-951-3 (1-1980)

QY 1 MetThrLysIlePheLysArgIleIleValPheAlaValIleAlaLeuSerSerGly 20  
Db 1 ATGACTACGAAATTTTAAAGGATCATTTGTTGCTGTATTTGCCCTATCGTCGGA 60  
QY 21 AsnIleLeuAlaGlnSerSerSerIleThrArgLysAspPheAspHisIleAsnLeuGlu 40  
Db 61 AATATACTGCACAAAGCTCTTCCATTACGAGGAAAAGATTTTGACCAACATCAACCTTGAG 120

QY 41 TyrSerGlyLeuGluLysValAsnLysAlaValAlaAlaGlyAsnTyrAspAspAlaAla 60  
Db 121 TATTCGGACTCGAAAAGGTTAATAAGCAGTTGCTCGCGCAACTATGACGATGCGGC 180  
QY 61 LysAlaLeuLeuAlaTyrTyrArgGluLysSerLysAlaArgGluProAspPheSerAsn 80  
Db 181 AAAGCATTTACTGGCATACTACAGGGAAGAAAGTAGGCCAGGAACTGATTTTCAGTAAT 240  
QY 81 AlaGluLysProAlaAspIleArgGlnProIleAspLysValThrArgGluMetAlaAsp 100  
Db 241 GCAGAAAAGCCTGCCGATATACGCCAGCCCATAGATAGTTACGGCTGAAATGGCCGAC 300  
QY 101 LysAlaLeuValHisGlnPheGlnProHisLysGlyTyrGlyTyrPheAspTyrGlyLys 120  
Db 301 AAGGCTTTTGGTCCACCAGTTTCAACCGCACAAAGGCTACGGCTATTTTTCATATGTTAA 360  
QY 121 AspIleAsnTyrGlnMetTyrProValLysAspAsnGluValArgTyrGlnLeuHisArg 140  
Db 361 GACATCACTGGCAGATGTCGGCGGTAAAGACAATGAAGTACGCTGGCAGTTGCCCGCT 420  
QY 141 ValLysTyrTyrGlnAlaMetAlaLeuValTyrHisAlaThrGlyAspGluLysTyrAla 160  
Db 421 GTAAAATGGTGGCAGCTATGCGCCCTGTTTATCAGCTACGGCGATGAAAATATGCA 480  
QY 161 ArgGluTyrValTyrGlnTyrSerAspTyrAlaArgLysAsnProLeuGlyLeuSerGln 180  
Db 481 AGAGAAATGGGTATATCAGTACAGCATTTGGGCGAGAAAACCCATTTGGCGCTGTCGAG 540  
QY 181 AspAsnAspLysPheValTyrArgProLeuGluValSerAspArgValGlnSerLeuPro 200  
Db 541 GATAATGATAAATTTGTGTCGGCGGCCCTTTGAAGTGTGCGACAGGTACAAAGTCTTCC 600  
QY 201 ProThrPheSerLeuPheValAsnSerProAlaPheThrProAlaPheLeuMetGluPhe 220  
Db 601 CCAACCTTCAGCTTATTTGTAAACTCGCCAGCCTTTTACCCAGCCTTTTAAATGAATTT 660  
QY 221 LeuAsnSerTyrHisGlnGlnAlaAspTyrLeuSerThrHisTyrAlaGluGlnGlyAsn 240  
Db 661 TTAACACAGTTACCCACCAACAGCCGATTTATTTATCTACGCATTATGCCGAAACAGGAAAC 720  
QY 241 HisArgLeuPheGluAlaGlnArgAsnLeuPheAlaGlyValSerPheProGluPheLys 260  
Db 721 CACCGTTTATTTGAAGCCCAACCACTTTGTTGAGGGGTATCTTTCCTCGAATTTAA 780  
QY 261 AspSerProArgTyrArgGlnThrGlyIleSerValLeuAsnThrGluIleLysLysGln 280  
Db 781 GATTCAACCAAGATGAGGCAACACCGCATATCGTGTGAAACACCGAGATCAAAAACAG 840  
QY 281 ValTyrAlaAspGlyMetGlnPheGluLeuSerProIleTyrHisValAlaAlaIleAsp 300  
Db 841 GTTTATGCCGATGGGATGCGATTTGAACCTTTCCACAAATTTACCATGTAGTCCCATCGAT 900  
QY 301 IlePheLeuLysAlaTyrGlySerAlaLysArgValAsnLeuGluLysGluPheProGln 320  
Db 901 ATCTTCTTAAAGCCCTATGGTTCTGCAAAACAGATTAACTTGAAAAGAAATTTCCGCA 960  
QY 321 SerTyrValGlnThrValGluAsnMetIleMetAlaLeuIleSerIleSerLeuProAsp 340  
Db 961 TCTTATGTACAACTGTAGAAAATATGATTATGCGGCTGATCAGTATTTTCACTGCGAGAT 1020  
QY 341 TyrAsnThrProMetPheGlyAspSerTyrIleThrAspLysAsnPheArgMetAlaGln 360  
Db 1021 TATAACACCCCTATGTTTGGAGATTCATGGATTACAGATAAAATTTTCAGGATGCGACAG 1080  
QY 361 PheAlaSerTyrAlaArgValPheProAlaAsnGlnAlaIleLysTyrPheAlaThrAsp 380  
Db 1081 TTTGCGAGCTGGCCCGGTTTTCGGCAACACAGGCCATATAAATATTTTGTACAGAT 1140  
QY 381 GlyLysGlnGlyLysAlaProAsnPheLeuSerLysAlaLeuSerAsnAlaGlyPheTyr 400  
Db 1141 GCAAAACAAAGGTAAGGCGCTTAACCTTTTATCCAAAGCATTTGAGCAATGACAGCTTTTAT 1200  
QY 401 ThrPheArgSerGlyTyrAspLysAsnAlaThrValMetValLeuLysAlaSerPro 420

```
Db 1201 AC GTT TAG CCG GAT GGG AT AAAAATG CAC CCGTATG GTATTA AAG CCG CTT CCCC 1260
Qy 421 Gly Glu Phe His Ala Gln Pro Asp Asn Gly Thr Phe Glu Leu Phe Ile Lys Gly Arg Asn 440
Db 1261 GGG GAATTT CAT GCG CCG GAT A A CCG GACTTTT GAAC TTTTAT AAG GCG CAG A A C 1320
Qy 441 Phe Thr Pro Asp Ala Gly Val Phe Val Tyr Ser Gly Asp Glu Ala Ile Met Lys Leu Arg 460
Db 1321 TTT ACC CCG ACG CCG GGT ATTT GTG TAT AG CCG CCG CAG CAG C C A T C A T G A A A C T C G G 1380
Qy 461 Asn Trp Tyr Arg Gln Thr Arg Ile His Ser Thr Leu Thr Leu Asp Asn Gln Asn Met Val 480
Db 1381 AACTGGTACCGTCAACCGCGATACAGACGCTTACCTCGACATCAAAATATGGTC 1440
Qy 481 Ile Thr Lys Ala Arg Gln Asn Lys Trp Glu Thr Gly Asn Asn Leu Asp Val Leu Thr Tyr 500
Db 1441 ATTACCAAGCCCGCAAAACAAATGGAAACAGGAAATACCTTGATGCTTACCTAT 1500
Qy 501 Thr Asn Pro Ser Tyr Pro Asn Leu Asp His Gln Arg Ser Val Leu Phe Ile Asn Lys Lys 520
Db 1501 ACCAACCACAGCTATCCGAATCTCGACCATCAGCGCAGTGTA CTTTTCATCAACAAAAA 1560
Qy 521 Tyr Phe Leu Val Ile Asp Arg Ala Ile Gly Glu Ala Thr Gly Asn Leu Gly Val His Trp 540
Db 1561 TACTTTCTGGTCATCGATAGGCGCATAGGCGAGCTACCGGAAACCTGGCGGTACACTGG 1620
Qy 541 Gln Leu Lys Glu Asp Ser Asn Pro Val Phe Asp Lys Thr Lys Asn Arg Val Tyr Thr 560
Db 1621 CAGCTTAAAGAAAGACAGCAACCCCTGTTTCGATAAGACAAAGAACCCGGGTTTACACCACT 1680
Qy 561 Tyr Arg Asp Gly Asn Asn Leu Met Ile Gln Ser Leu Asn Ala Asp Arg Thr Ser Leu Asn 580
Db 1681 TACAGAGATGTATCAACCTGATCATCAATCGTTGAATGCGACAGGACCGACCTCAAT 1740
Qy 581 Glu Glu Glu Lys Val Ser Tyr Val Tyr Asn Lys Glu Leu Lys Arg Pro Ala Phe Val 600
Db 1741 GAAGAGAGAAAGGATCTTATGTTCATTAAGGAGCTGAAAGACCTGCTTTCGTA 1800
Qy 601 Phe Glu Lys Pro Lys Lys Asn Ala Gly Thr Gln Asn Phe Val Ser Ile Val Tyr Pro Tyr 620
Db 1801 TTTCAAAGCCATAAAGAAATGCGGCACACAAAAATTTGTGTCAGTAGATTTATCCATAC 1860
Qy 621 Asp Gly Gln Lys Ala Pro Glu Ile Ser Ile Arg Glu Asn Lys Gly Asn Asp Phe Glu Lys 640
Db 1861 GACGGCAGAGAGCTCCAGAGATCAGCATACGGGAAACAAAGGCAATGATTTTGAGAA 1920
Qy 641 Gly Lys Leu Asn Leu Thr Thr Ile Asn Gly Lys Gln Gln Leu Val Leu Val Pro 659
Db 1921 GGCAAGCTTAATCTAACCCCTTACCATTAAACGAAACAAACAGCTTGTGTTGTTCT 1977
```

## RESULT 3

```
PCT-US95-07391A-3
; Sequence 3, Application PC/TUS9507391A
; GENERAL INFORMATION:
; APPLICANT: IBEX TECHNOLOGIES and
; APPLICANT: ZIMMERMANN, Joseph
; TITLE OF INVENTION: Nucleic Acid Sequences And Expression
; TITLE OF INVENTION: Systems For Heparinase II And Heparinase III Derived From
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hale and Dorr
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: PCT/US95/07391A
; FILING DATE: 09-JUNE-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/258,639
; FILING DATE: 10 JUNE 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: BAKER, Hollie L.
; REGISTRATION NUMBER: 31,321
; REFERENCE/DOCKET NUMBER: 104385.116PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)942-8400
; TELEFAX: (202)942-8484
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1980 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; PCT-US95-07391A-3
;
; Alignment Scores:
; Pred. No.: 0 Length: 1980
; Score: 3494.00 Matches: 659
; Percent Similarity: 100.00% Conservative: 0
; Best local Similarity: 100.00% Mismatches: 0
; Query Match: 100.00% Indels: 0
; DB: 5 Gaps: 0
;
; US-09-802-285A-2 (1-659) x PCT-US95-07391A-3 (1-1980)
```

```
Qy 1 Met Thr Thr Lys Ile Phe Lys Arg Ile Ile Val Phe Ala Val Ile Ala Leu Ser Ser Gly 20
Db 1 ATGCTTACGAAAATTTTTTAAAGAGATCATTTGTTCTGTTAAATGGCTTATCGTCGGGA 60
Qy 21 Asn Ile Leu Ala Gln Ser Ser Ile Thr Arg Lys Asp Phe Asp His Ile Asn Leu Glu 40
Db 61 AATATATTGCAACAAGCTCTTCCATTACAGGAAAGATTTTGACCATCACTCACTTGGAG 120
Qy 41 Tyr Ser Gly Leu Glu Lys Val Asn Lys Ala Val Ala Ala Gly Asn Tyr Asp Asn Ala Ala 60
Db 121 TATTCCGACTCGGAAAAGGTTAATAAAGCAGTTCTGTCGCGCAACTATGACGATGCGGCC 180
Qy 61 Lys Ala Leu Leu Ala Tyr Tyr Arg Glu Lys Ser Lys Ala Arg Glu Pro Asp Phe Ser Asn 80
Db 181 AAACATTTACTGGCATACTACAGGGAAGAAAGTAAGCCGGAACCTGATTTTCAGTAAT 240
Qy 81 Ala Glu Lys Pro Ala Asp Ile Arg Gln Pro Ile Asp Lys Val Thr Arg Glu Met Ala Asp 100
Db 241 GCAGAAAAGCCCTCGCATATACGCCAGCCCATAGATAAGGTTACGCGTGAATGCGGCAC 300
Qy 101 Lys Ala Leu Val His Gln Phe Gln Pro His Lys Gly Tyr Gly Tyr Phe Asp Tyr Gly Lys 120
Db 301 AAGCGTTTGTGTCACCAAGTTTCAACCGCACAAAGGCTACGCGCTATTTTGATTATGTTGTA 360
Qy 121 Asp Ile Asn Trp Gln Met Trp Pro Val Lys Asp Asn Glu Val Arg Trp Gln Leu His Arg 140
Db 361 GACATCAACTGGCAGATGTGCCGTTAAAGACAAATGAAGTACCTGGCAGTTGACCGGT 420
Qy 141 Val Lys Trp Trp Gln Ala Met Ala Leu Val Tyr His Ala Thr Gly Asp Glu Lys Tyr Ala 160
Db 421 GTAAAAATGGTCGAGGCTATGGCCCTGTTTATCACGCTACCGGCGATGAAAAATATGCA 480
Qy 161 Arg Glu Trp Val Tyr Gln Tyr Ser Asp Trp Ala Arg Lys Asn Pro Leu Gly Leu Ser Gln 180
Db 481 AGAAGATGGGTATATCATAGTACAGCGATTGGGCCAGAAAAACCATTTGGGCTGTGCGAG 540
Qy 181 Asp Asn Asp Lys Phe Val Trp Arg Pro Leu Glu Val Ser Asp Arg Val Gln Ser Leu Pro 200
Db 541 GATATATGATAAATTTGTGTGCGGCCCTTCAAGTGTGCGACAGGTTACAAAGTCTTCCC 600
Qy 201 Pro Thr Phe Ser Leu Phe Val Asn Ser Pro Ala Phe Thr Pro Ala Phe Leu Met Glu Phe 220
```

Db 601 CCAACCTTCAGCTTATTTGTAACCTGCGCAGCTTTTACCCCGACCTTTTAAATGGAATTT 660  
QY 221 LeuAsnSerTyrHisGlnGlnAlaAspTyrLeuSerThrHisTyrAlaGluGlnGlyAsn 240  
Db 661 TTAACAGTTACACCAACAGCGCGGNTATTTATCTACGCATTTAGCCGAACAGGGAAC 720  
QY 241 HisArgLeuPheGluAlaGlnArgAsnLeuPheAlaGlyValSerPheProGluPheLys 260  
Db 721 CACCGCTTATTTGAAGCCCAACGCACTTGTGTCAGGGGTATCTTTCCCTGAAATTTAA 780  
QY 261 AspSerProArgTyrArgGlnThrGlyIleSerValLeuAsnThrGluIleLysLysGln 280  
Db 781 GATTCACCAAGATGAGGCAACCGGCATATCGGTCTGAACACCGAGATCAAAACACAG 840  
QY 281 ValTyrAlaAspGlyMetGlnPheGluLeuSerProIleTyrHisValAlaAlaIleAsp 300  
Db 841 GTTTATGCCGATGGATGAGTTTGAACCTTTTACCAATTTTACCATGTAGCTGCCATCGAT 900  
QY 301 IlePheLeuLysAlaTyrGlySerAlaLysArgValAsnLeuGluLysGluPheProGln 320  
Db 901 ATCTTCTTAAGCGCTATGTTCTGCAAAACGAGTTAACTTTGAAACAAATTTTCGCCAA 960  
QY 321 SerTyrValGlnThrValGluAsnMetIleMetAlaLeuIleSerIleSerLeuProAsp 340  
Db 961 TCTTATGTACAACTGTAGAAATATGATTATGGCGCTGATCAGTATTTTCACTGCCAGAT 1020  
QY 341 TyrAsnThrProMetPheGlyAspSerTrpIleThrAspLysAsnPheArgMetAlaGln 360  
Db 1021 TATAACACCCCTATGTTTGGAGATTCATGATACAGATAAAAAATTTCAAGGATGCCACAG 1080  
QY 361 PheAlaSerTrpAlaArgValPheProAlaAsnGlnAlaIleLysTyrPheAlaThrAsp 380  
Db 1081 TTTTGCAGCTGGCGCGGTTTTCGGGCAACACGAGCCCATATAAATTTTGTCTACAGAT 1140  
QY 381 GlyLysGlnGlyLysAlaProAsnPheLeuSerLysAlaLeuSerAsnAlaGlyPheTyr 400  
Db 1141 GGCNAACAGGTAGCGCGCTAACTTTTATCCAAAGCAATTGAGCAATGACGCTTTTAT 1200  
QY 401 ThrPheArgSerGlyTrpAspLysAsnAlaThrValMetValLeuLysAlaSerProPro 420  
Db 1201 ACGTTTAGAGCGGATGGGATAAAATGCAACCGTTATGGTATTAAGACCGAGTCTCTCC 1260  
QY 421 GlyGluPheHisAlaGlnProAspAsnGlyThrPheGluLeuPheIleLysGlyArgAsn 440  
Db 1261 GGGGAATTTTCATGCCAGCGGATACGGGACTTTTGAACCTTTTATAAAGGCGCAGAAAC 1320  
QY 441 PheThrProAspAlaGlyValPheValTyrSerGlyAspGluAlaIleMetLysLeuArg 460  
Db 1321 TTTACCCAGACGCGCGGTTATTTGTATAGCGGCGAGCAAGCCATCATGAACTGCCG 1380  
QY 461 AsnTrpTyrArgGlnThrArgIleHisSerThrLeuThrLeuAspAsnGlnAsnMetVal 480  
Db 1381 AACTGTGACCGTCAAAACCCGCATACACAGCAGCTTACACTCGACAATCAAAATATGTC 1440  
QY 481 IleThrLysAlaArgGlnAsnLysTrpGluThrGlyAsnAsnLeuAspValLeuThrTyr 500  
Db 1441 ATTACAAAGCCCGCAACAAATGGGAAACAGGAAATAAATCTGATGTGCTTACCTAT 1500  
QY 501 ThrAsnProSerTyrProAsnLeuAspHisGlnArgSerValLeuPheIleAsnLysLys 520  
Db 1501 ACCAACCCAGCTATCCGAACTTGGACCATCAGCGAGTGTACTTTTCATCAACAAAAA 1560  
QY 521 TyrPheLeuValIleAspArgAlaIleGlyGluAlaThrGlyAsnLeuGlyValHisTrp 540  
Db 1561 TACTTCTGCTCATCGATAGGCAATAGCGAAGCTACCGGAAACCTGGCGGTACACTCG 1620  
QY 541 GlnLeuLysGluAspSerAsnProValPheAspLysThrLysAsnArgValTyrThrThr 560  
Db 1621 CAGCTTAAAGAGACAGCAACCTGTTTTCGATTAAGCAAAAGAACCGGGTTTACACCAT 1680  
QY 561 TyrArgAspGlyAsnAsnLeuMetIleGlnSerLeuAsnAlaAspArgThrSerLeuAsn 580

Db 1681 TACAGAGATGGTAACAACCTGATGATGATGATGATGATGATGATGATGATGATGAT 1740  
QY 581 GluGluGluGlyLysValSerTyrValTyrAsnLysGluLeuLysArgProAlaPheVal 600  
Db 1741 GAAGAAGAAGGAAGGATATCTTATGTTTACATAAAGGAGCTGAAAAAGACCTGCTTCGTA 1800  
QY 601 PheGluLysProLysLysAsnAlaGlyThrGlnAsnPheValSerIleValTyrProTyr 620  
Db 1801 TTTGAAAAAGCTTAAAAAAGAAATCCGGCACACAAAAATTTTGTGATATAGTTTATCCATAC 1860  
QY 621 AspGlyGlnLysAlaProGluIleSerIleArgGluAsnLysGlyAsnAspPheGluLys 640  
Db 1861 GACGGCCAGAGCTCCAGAGATCAGCATACGGGAAACAAAGGCAATGATTTTGAAGAA 1920  
QY 641 GlyLysLeuAsnLeuThrLeuThrIleAsnGlyLysGlnGlnLeuValLeuValPro 659  
Db 1921 GGCAAGCTTAATCTTAACCCCTTACCATTACCGAAACCAACAGCTTGTGTGTTCTCT 1977

## RESULT 4

US-08-961-527-186/c  
; Sequence 186, Application US/08961527  
; Patent No. 6420135  
; GENERAL INFORMATION:  
; APPLICANT: Charles Kunsch  
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences  
; NUMBER OF SEQUENCES: 391  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,527  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340PI  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 186:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3763 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-961-527-186

## Alignment Scores:

Pred. No.: 9,81e-10 Length: 3763  
Score: 176.00 Matches: 116  
Percent Similarity: 34.63% Conservative: 80  
Best Local Similarity: 20.49% Mismatches: 231  
Query Match: 5.04% Indels: 140  
DB: 4 Gaps: 20

US-09-802-285A-2 (1-659) x US-08-961-527-186 (1-3763)

QY 71 SerLysAlaArgGluProAspPheSerAsnAlaGluLysProAlaAspIleArgGlnPro 90

Db 2173 TCAAAAGACTATGAGAAAGTCAAGAGTTCCCTTTGAACGCTTGATGACAATCGTTTATG 2114



Db 208 AAAGAGTATCGCGTGAATCGGTATCCTGATGATGATCCGAATGGCTCTTTATGTTGAGT 267  
Qy 140 ArgValValThrPrlGlnAlaMetAlaLeuValTyrHisAlaThrGlyAspGluTyr 159  
Db 268 CGTCAAAAGCTTTCTTGTAGATCTGGCACAAGCATATGCACTTACTAAAAAGAACGTTAC 327  
Qy 160 AlaArgGluThrValTyrGlnTyrSerAspTrpAlaArgLysAsnProLeuGlyLeuSer 179  
Db 328 TTACAGAAATGGCAGACTTGTCTTATTTGATTTTATTT-----AACGATGAGGTGAGCCA 381  
Qy 180 GlnAspAsnAspLysPheValTyrArgProLeuGluValSerAspArgValGlnSerLeu 199  
Db 382 AATTCACGATAGGATGTTGGCGTCCGTAGATGTTGGGATTCGATTAACAAC-----438  
Qy 200 ProProThrPheSerLeuPheValAsnSerProAlaPheThrProAlaPheLeuMetGlu 219  
Db 439 -----TGGATGAAAGCTTCGACGTATATTCCAAATCGCTGATTTTCAGA 480  
Qy 220 PheLeu-----AnSerTyrHisGlnGlnAlaAspTyrLeuSer 232  
Db 481 CTATTAGGAATGATGATGTTGAAACAGCGCTTCGTCATCCATCTGGCATTTTGGAG 540  
Qy 233 ThrHisTyrAlaGluGlnGlyAsnHisArgLeuPheGlu-----245  
Db 541 CGTCTCTATTCGATPAAA-----TACAGCTCAGTAATTTGGGCTGTGTTGGCAATGGT 594  
Qy 246 -----AlaGlnArgAsnLeuPheAlaGlyValSerPheProGluPheLysAspSerPro 263  
Db 595 GGAATGGCAGCTATTGATTTATTT-----CTTCCAGAACTGGTGACCACT---639  
Qy 264 ArgTrpArgGlnThrGlyLeuSerValLeuAsnThrGluLeuLysGlnValTyrAla 283  
Db 640 AAACAGAGGATCAATATATGTTCTGCTGTAGCTGAACAGCTGATCAACAATTTCTATTCA 699  
Qy 284 AspGlyMetGlnPheGluLeuSerProIleTyrHisValAlaAlaIleAspIlePheLeu 303  
Db 700 GATGGATTCATTTGGAGCAGAGCCCGCTGTAC-----732  
Qy 304 LysAlaTyrGlySerAlaLysArgValAsnLeuGluLysGluPheProGlnSerTyrVal 323  
Db 733 -----CAGCAGCAAGTTTGTGATGACATTCGTG 759  
Qy 324 GlnThrValGluAsnMetIleMetAlaLeuIleSerIleSerLeuPro-----339  
Db 760 TATCTATTCGAG-----ATTCTGAATATCTTGAAGTCGAGTACCATAGATCTTCGC 813  
Qy 340 ---AspTyrAsnThrProMetPheGlyAspSerTrpIleThrAspLys-----354  
Db 814 ATGAAATTAACACCTATTCTCTACCCATATTTCGCGGATAACCAAGATATCCTTA 873  
Qy 355 -----AsnPhe-----356  
Db 874 AATCCGATCAATGATGATCATGTCACCTTTTCATTAGCTATGATATCTATTCGCAAA 933  
Qy 357 -----ArgMetAlaGlnPheAlaSerTrpAla 365  
Db 934 TTAGGTTTCATATTGAACCTCCATGACTGCGAATATGCAAGCTT-----TGGAGC 987  
Qy 366 -----ArgValPheProAlaAsnGlnAlaIleLysTyrPheAla---378  
Db 988 GGGGATCTTTATGAAGAAAGATCTGGGAAAC-AATGAAGCCAAAAGAACTTTTTCGTGG 1046  
Qy 379 -----ThrAspGlyLysGlnGlyLeuAlaProAsnPheLeuSerLysAlaLeu 394  
Db 1047 CGAATCAAGTGGGCTGATGCGGTACAAAGCAGA-----GGATATCTTA 1088  
Qy 395 SerAsnAlaGlyPheTyrThrPhe---ArgSerGlyTrpAspLysAsnAlaThrValMet 413  
Db 1089 -----TTTATACCTTTTAAACGCTCTGCATGG-----1115  
Qy 414 ValLeuLysAlaSerProGlyGlu-PheHisAlaGlnProAspAsnGlyThrPheGlu 433  
Db 1116 -----GAGCGCACATGGTCATGCTTCTACAGTGATTTAC 1151

Qy 433 uLeuPheIleLysGlyArgAsnPheThrProAspAlaGlyValPheValTyrSerGlyAs 453  
Db 1152 ACTCAACATACAAAGGGATGACTTATTTCCGATAGTGTGCTTACAGCTATGTCAACA 1211  
Qy 453 pGluAlaIleMetLysLeuArgAsnTrpTyrArgGlnThrArgIleHisSerThrLeuTh 473  
Db 1212 ATCAGAGCGACTTCAGCTA-----AAAGAGTGGCTTCGCAATACGATGTT 1259  
Qy 473 rLeu---AspAsnGlnAsnMetValIleThr-----LysAlaArgGlu 486  
Db 1260 TATCGCGAATAATCCCATACCTTATTTCCGATAGTGGGTATGACAAATTACCGAC 1319  
Qy 486 nAsnLysTrpGluThrGlyAsnLeuAspValLeuThrTyrThrAsnProSerTyrPr 506  
Db 1320 ACCCTTATTTCCAGAAATAAAGAACTTCTGTGCGTTTTTTTTCGAGAATGCGCTGGCT 1379  
Qy 506 oAsnLeuAspHisGlnArgSerValLeuPheIleAsnLysLysTyrPheLeu-----523  
Db 1380 GGATAAGCGGATCCAGATCCCAATGATTTTTCGAGCGCAGCTTCATCTATTTAAAGTCGAT 1439  
Qy 524 -----ValIleAspArgAlaIleGlyGlu-----AlaTh 533  
Db 1440 CAACTCGGTAGTTATTTATGATAGCTTTCGAGGACAGAAAGACTGAAATTACGAGTAC 1499  
Qy 533 rGlyAsnLeuGlyValHisTrpGlnLeuLysGluAspSerAsnProValPheAspLysTh 553  
Db 1500 CTATATTTGGCAGCGCTCGATTAATTTGTCAAAAGAAAGCGCATCGGTTTGTCTTAACACTAC 1559  
Qy 553 rLysAsnArgValTyrThrTyrArgAspGlyAsnAsnLeuMetIleGlnSerLeuAs 573  
Db 1560 GAACAGCATAAGTACACACTCTCTTTTCTCTGGAGGC-----1596  
Qy 573 nAlaAspArgThrSerLeuAsnGluGluGlyLysValSerTyrValTyrAsnLysGlu 593  
Db 1597 -----CAAACACAGCAATCAGTGGCAAAAGGTTTCAGAGATTTATAACCACT 1643  
Qy 593 uLeuLysArgPro-----AlaPheValPheGluLysProLysLysAsnAlaGlu 609  
Db 1644 GAATGAGCACCACCGCTTATCAACAAGTTTGTTCACAAACGGG-AAAAGAAATTCAAG 1702  
Qy 609 YThrGlnAsnPheValSerIleValTyrProTyrAspGlyGlnLysAla 625  
Db 1703 CAAAGCTCAITTCCTCGTTAGAGGATATCCAAATTACGCCAATAAAGT 1751

## RESULT 6

US-09-134-001C-795  
; Sequence 795, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; PRIOR FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 795  
; LENGTH: 2097  
; TYPE: DNA  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-795

Alignment Scores: 4,46e-05 Length: 2097  
Pred. No.: 132.50 Matches: 140  
Score: 34.59% Conservative: 108  
Best Local Similarity: 19.53% Mismatches: 260  
Query Match: 3.79% Indels: 210

```
DB: 4 Gaps: 30
US-09-802-285A-2 (1-659) x US-09-134-001C-795 (1-2097)
Qy 9 IleIleValPheAlaValIleAlaLeuSerSerGlyAsnIleLeuAlaGlnSerSer 27
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 94 GTAATTATCTTTGCTATAGTCGTATGAGATTAGATTATTTACAAATAGCACAAGGATCT 153
Qy 28 SerIleThrArg-----LysAspPheAspHisIleLeuLeuGluTyrSerGlyLeuGlu 45
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 154 CATTACAAACAATAATCAAAAACGATGAAAACATAACT----- 192
Qy 46 LysValAsnLysAlaValAlaAlaGlyAsn-----Tyr 56
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 193 ---GTTAATGATCAGTACCAAGAGCGCGAATACTAGATAGAAATGCGAAGACTACTGTT 249
Qy 57 AspAspAlaAlaLysAlaLeuLeuAlaTyrTyrArgGluLysSerLysAlaArgGluPro 76
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 250 GATAATGCTTCAAGATGCTATTATACATACACTAGAACCCGTAAAAACATCACAAAAGGAA 309
Qy 77 AspPheSerAsnAlaGluLysProAlaAsp---IleArgGlnProIleAspLysValThr 95
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 310 ATGTAAATACTGCTAGAAATCTCACAGATTAAATTAATAATGCGATACAGATAAAATTA 369
Qy 96 ArgGluMetAlaAspLysAlaLeuValHisGlnPheGlnProHisLysGlyTyrGlyTyr 115
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 370 GAGAGA-----GATAAA----- 381
Qy 116 PheAspTyrGlyLysAspIleAsnTrpGlnMetTrpPro-----ValLysAspAsn 132
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 382 -----AAGGATTTTGGATTCAAAATGATCGTCACTTTGCTAAAAAGTT-AA 428
Qy 133 GluValArgTrpGlnLeuHisArgValLysTrpTrpGln----- 145
Db 429 GAGAAAAGA-----ACAATTAAATGTTAGAGGTGGCAGTATTTCAACAAGCAATTTGA 482
Qy 146 ----- 147
Db 483 TACCCAACTTAGATATAAATAGGAAAAAACAATTAACAGATTAACTAAAAAGATT 542
Qy 148 AlaLeuValTyrHisAlaThrGlyAspGluLysTyrAla----- 160
Db 543 GCAAGTTTTTAGCAATTTATCGGAAATGAACGCTGGTCACTCTAGATCTCCTCAACAAT 602
Qy 161 -----ArgGlu----- 162
Db 603 TAAAAATGAACGACGTACCGAAGAAAGAAATATGCGAGCGGTATCAACAGCTTTCTAAATT 662
Qy 163 ---TrpValTyrGlnTyrSer-AspTrpAlaArgLysAsnProLeuGlyLeuSerGlnAs 181
Db 663 ACCTGGTGTTAAATACTCAATGGATTGGATAGAAAATACCCTACGCT-----GA 713
Qy 181 pAsnAspLysPheValTrpArgProLeuGluValSerAspArgValGlnSerLeuProPr 201
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 714 TACTTTAAGAGGCATATTGGA-----GATGTGTCGACTTCGACTGAAGGTATACCTTAA 767
Qy 201 oThrPheSerLeuPheValAsnSerProAlaPheThrProAlaPheLeuMetGluPheLe 221
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 768 AGAATTAAT----- 788
Qy 221 uAsnSerTyrHisGlnGlnAlaAspTyrLeuSerThrHisTyrAlaGluGlnGlyAsnHi 241
Db 789 ATCAAAAGGTTTATTCAGAAATGATCGGTCGGTAAATCTTATCTTGAATATCAATACGA 848
Qy 241 sArgLeuPheGluAlaGlnArgAsnLeuPheAlaGlyValSerPheProGluPheLysAs 261
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 849 AGATGTTCTTAAGGCCAGCAAGAAACAATG-----AAATATACAC 890
Qy 261 pSerProArgTrpArgGlnThrGlyIleSerValLeuAsnThrGluIleLysGlnVa 281
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 891 TGATAAATCTCGAAGAGTAATAAGTTCAAGAGTACTTAATCTCGCTCAAGA----- 942
Qy 281 lTyrAlaAspGlyMetGlnPheGluLeuSerProIleTyrHisValAlaAlaIleAspIl 301
```

Db 1887 CGAGAACTAGATTAACTCAACAT---ATATCGTTATGACCCG 1930

# RESULT 7

US-07-973-320-1  
 ; Sequence 1, Application US/07973320  
 ; Patent No. 5286486  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Payne, Jewel M.  
 ; APPLICANT: Fu, Jenny M.  
 ; TITLE OF INVENTION: No. 5286486el Bacillus thuringiensis Gene  
 ; TITLE OF INVENTION: Encoding a Coleopteran-Active Toxin  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: David R. Saliwanchik  
 ; STREET: 2421 N.W. 41st Street, Suite A-1  
 ; CITY: Gainesville  
 ; STATE: FL  
 ; COUNTRY: USA  
 ; ZIP: 32606  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent In Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/973,320  
 ; FILING DATE: 19921106  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/788,638  
 ; FILING DATE: 6-NOV-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Saliwanchik, David R.  
 ; REGISTRATION NUMBER: 31,794  
 ; REFERENCE/DOCKET NUMBER: MA68.C1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 904-375-8100  
 ; TELEFAX: 904-372-5800  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3414 base pairs  
 ; TYPE: NUCLEIC ACID  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; HYPOTHETICAL: NO  
 ; ANTI-SENSE: NO  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Bacillus thuringiensis  
 ; STRAIN: dakota  
 ; INDIVIDUAL ISOLATE: HD511  
 ; IMMEDIATE SOURCE:  
 ; LIBRARY: Lamdagem (TM)-11 library of J.M. Fu  
 ; CLONE: 511  
 ; US-07-973-320-1

## Alignment Scores:

Pred. No.: 0.0011 Length: 3414  
 Score: 123.50 Matches: 141  
 Percent Similarity: 35.92% Conservative: 95  
 Best Local Similarity: 21.46% Mismatches: 199  
 Query Match: 3.53% Indels: 224  
 DB: 1 Gaps: 39

US-09-802-285A-2 (1-659) x US-07-973-320-1 (1-3414)

QY 48 AsnLysAlaValAla-----AlaGlyAsnTyrAspAlaAlaLysAlaLeu 63  
 Db 361 AATAAGAGCATTCGAGAAATAGAGGATTTAGGAAATAAC-----399  
 QY 64 LeuAlaTyrArgGluLysSerLysAlaArgGluProAspPheSerAsnAlaGluLys 83  
 Db 400 TTAACAATATAACACAG-----GCATTGAAGATTGGCTGAACAATCCTGATGAT 450

QY 84 ProAlaAspIleArgGlnProIleAspLysValThrArgGluMetAlaAspLysAlaLeu 103  
 Db 451 CCAGCACTATAACACAGAGTAGATCGTTTT-----CGTATATTAGAT---GCTTTA 501  
 QY 104 ValHisGlnPheGlnProHisLysGlyTyrGlyTyrPheAsp-----117  
 Db 502 TTTGAATCATATATGCCCTCATTTAGGGTTGCTGGATATGAAATACCATTAACAGTT 561  
 QY 118 TyrGlyLysAspIleAsnTrpGlnMet-----126  
 Db 562 TAGCACAAGCGGCAACCTTCATCTAGCTTTATTAAGAGATTCTACTCTTTATGAGAT 621  
 QY 127 ---TrpProValLysAspAsnGluVal-----134  
 Db 622 AAATGGGATTTCACATCAGAACCAACATTCAGGAAAAATTAAATCGTCAAAAGAAACATATC 681  
 QY 135 ---ArgTrpGlnLeuHisArgValLysTrpTrpGlnAla---MetAlaLeuValTyrHis 152  
 Db 682 TCTGAATATTCTAACCATTCGTTAAGTGTATATAGTGTCTTAGCAGATTGAACGGT 741  
 QY 153 AlaThrGlyAspGlu-----LysTyrAlaArgGluTrpValTyrGlnTyr 167  
 Db 742 TCCACTTATGAACAATGGATAAATTATAATCGTTTTCTGAGAGAAATGATATTAAATGGTA 801  
 QY 168 SerAspTrpAlaArgLysAsnProLeu-----GlyLeu 178  
 Db 802 TTAGATATTGCTGCTGTTATTCCTATTATGACCCTCGAATGTATTCAATGAAACAACT 861  
 QY 179 SerGlnAspAsnAspLysPheValTrpArg-ProLeuGluValSerAspArgValGlnSe 198  
 Db 862 AGCAGTT-AACGAGAGAGTGTATACCATTCATAGCTTGTCAATTAGCAATCCAGA 920  
 QY 198 rLeuProThrPheSerLeuPheValAsnSerProAlaPheThrProAlaPheLeuMe 218  
 Db 921 TATAGTCCAGCTTTTCTCAGATGGAATAATCTCGTTTAGAACACACACCTTGT-- 978  
 QY 218 cGluPheLeuAsnSerTyrHisGlnAlaAspTyrLeuSerThrHisTyrAlaGluL 238  
 Db 979 -----GATTATTAGATGAGCTTTATATATATATAC 1007  
 QY 238 nGlyAsnHisArgLeuPhe-----GluAlaGlnArgAsnLeuPhe-----Al 252  
 Db 1008 ATCAAAATATAAGCATTTTCATGAGATTCAACAGACCTATTTTATTGTTGTGTACA 1067  
 QY 252 aGlyValSerPheProGluPheLysAspSerProArgTrpArgGlnThrGlyLe-SerV 272  
 Db 1068 TAAGGTTAGCTTTAAAAAATCGAGCAATCCCAATTATATAT---ACAACAGGCATATATCG 1124  
 QY 272 alLeuAsnThrGluLe-----LysLysGlnValTyrAlaAspGlyMet---GlnPheG 289  
 Db 1125 TAAACAAGTGGATATATTTTCATCAGGAGCATATTTCATTTAGAGGAATGATATCTATAG 1184  
 QY 289 luLeuSerProIleTyrHisValAlaAlaIleAspIlePheLeuLysAlaTyrGlySerA 309  
 Db 1185 AACATTAGCAGCTCCATCAGTTGTAGTTATCCGTATA---CTCAGAATTATGTTGTGCA 1241  
 QY 309 laLys-----310  
 Db 1242 GCAAGTTGAGTTTACGGTGTAAAAGGGCATGTACATTATAGAGGAGATAACAAATATGA 1301  
 QY 311 ----ArgValAsnLeuGluLysGluPheProGlnSerTyrValGlnThrValGluAsnM 329  
 Db 1302 TCTGAGGTATGATTCTTATGATCAATATACCCCGAGAGAACCAATACACGAAAAATA 1361  
 QY 329 etIleMetAlaLeu-----IleSerIleSerLeuProAspTyrAsnThr- 343  
 Db 1362 CACTCATCGA-TTATGTCTGCTACAGCTATATCTAAATCAACTCGGATTATGATAATG 1420  
 QY 344 -----ProMetPheGlyAspSerTrpIleThrAspLysAsnPheArgMetAlaGlnP 361  
 Db 1421 CTACTATCCGATCTTT-----TCTTGG-----ACGCATAGAGTGGGAGT 1462

```
Qy 361 heaLaSerTrpAlaArgValPheProAlaasn-----GlnAlaIleLysTyrP 377
Db 1463 AT-----TACATAGATCTATCCAAACAAATCAAAATTCACAGCTGTAAATAATGT 1516
Qy 377 heaLaThrAspGlyLysGlnGly-----LysAlaProAsnPhelSerLysAlaLeuS 395
Db 1517 ATAACTAGATGATCTATCTACAGTTGTCAAAGGCGCTGGATTACAGGTGAGATTAG 1576
Qy 395 erAsnAlaGlyPheTyrThrPheArgSerGlyTrp-----AsPlyAsnAlaThrValm 413
Db 1577 TTAAGAGAGGG-----AGTAATGTTTATATAGGAGATATAAAGGCTACCGTA- 1623
Qy 413 etValLeuLysAlaSerProProGlyGlu-----PheHisAlaGlnP 427
Db 1624 -----AATCACCATTCTCTCAAAATATCGTGTAGAGTTCGATACCCACTA 1672
Qy 427 roAspAsnGlyThrPheGluLeuPheIleLysGlyArgAsnPhelThrProAspAlaGlyV 447
Db 1673 GTGTTTCTGGACTATTCAACGTGTTTATT-----1701
Qy 447 alPheValTyrSerGlyAspGluAlaIleMetLysLeuArgAsnTrpTyrArgGlnThrA 467
Db 1701 -----1701
Qy 467 rgIleHisSerThrLeuThrLeuAspAsnGlnAsnMetValIleThrLysAlaArgGlnA 487
Db 1702 -----AATGATGAATAGCGCTTCAAAAATTTTCAA 1735
Qy 487 snLysTrpGluThr-----GlyAsnAsnLeuAspValLeuThrTyr-----500
Db 1736 GTACTGTAGAAACAATAGGTGAAGAAAGAT-----TTAACCTATGTTTCATTGGATATA 1792
Qy 501 -----ThrAsnProSerTyrProAsn-----LeuA 509
Db 1793 TAGAATATCTGACACCATTCATTTCCGAATGACATCCAAAATCCTCTTCATTAA 1852
Qy 509 sphis-----GlnArgSerValLeuPheIleAsnLysLysTyrPheLeuValIleAspA 527
Db 1853 ACCATTGAGTAACAATTCACCATTTATGTAGATTCATCGAATTTATCCCTGTAGAT- 1911
Qy 527 rgAlaIleGlyGluAlaThrGlyAsnLeuGlyValHisTrpGlnLeuLysGluAspSerA 547
Db 1912 -----GTAAATATTATGATGAAAAAGAAAAA-----1935
Qy 547 snProValPheAspLysThrLysAsnArgValTyrThrTyrArgAspGlyAsnAsnL 567
Db 1936 -----CTAGAAAAAGCACAGAAAGCCGGTAATACCTTGTTTTACAGAGGAGAAATG 1987
Qy 567 euMet-----IleGlnSerLeuAsnAlaAspArgThrSerLeu 579
Db 1988 CACTCCAAAATAGTGCACAGATTATATAAGTGGACCATGTTTCAATT 2034
```

## RESULT 8

```
US-08-328-962-1
; Sequence 1, Application US/08328962
; Patent No. 5536637
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; TITLE OF INVENTION: A NOVEL METHOD FOR ISOLATING CYTOKINES
; TITLE OF INVENTION: AND OTHER SECRETED PROTEINS
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Legal Affairs, Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/328,962
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/045,267
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcdaniels, Patricia A.
; REGISTRATION NUMBER: 33,194
; REFERENCE/DOCKET NUMBER: GI 5200
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-876-1170
; TELEFAX: 617-876-5851
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1542 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Saccharomyces cerevisiae
; IMMEDIATE SOURCE:
; CLONE: SUC2
; US-08-328-962-1
Alignment Scores:
Pred. No.: 0.000346 Length: 1542
Score: 123.00 Matches: 109
Percent Similarity: 32.01% Conservative: 76
Best Local Similarity: 18.86% Mismatches: 179
Query Match: 3.52% Indels: 214
DB: 1 Gaps: 31
US-09-802-285a-2 (1-659) x US-08-328-962-1 (1-1542)
```

```
Qy 94 ValThrArgGluMetAlaAspLysAlaLeuValHisGlnPheGlnProHisLysGlyTyr 113
Db 4 ATGACAAACGAACTAGCATAGACCTTTGGTCCAC---TTCCACCCCAACACAGGGCTGG 60
Qy 114 -----GlyTyrPheAspTyrGlyLysAspIleAsnTrpGlnMetTrp---127
Db 61 ATGATGACCAACAAATGGTTGTGGTAGCATGAAAGATGCCAATGGCATCTGACTTT 120
Qy 128 ProValLysAspAsnGluValArgTrpGlnLeuHisArgValLysTrpTrpGlnAlaMet 147
Db 121 CAATACAAACCAAAATGACACCGTATGG-----GGTACGCCATTGTTTGGGGC-----168
Qy 148 AlaLeuValTyrHisAlaThrGlyAspGluLysTyrAlaArgGluTrpValTyrGlnTyr 167
Db 169 -----CATGCTACTTCCGATGAT-----TTG 189
Qy 168 SerAspTrpAlaArgLysAsnProLeuGlyLeuSerGlnAspAsnAspLysPheValTrp 187
Db 190 ACTAATGG---GAAGATCAACCACTATGCTCTCCCAAGCGTAAC-----234
Qy 188 ArgProLeuGluValSerAspArgValGlnSerLeuProThrPheSerLeuPheVal 207
Db 234 -----234
Qy 208 AsnSerProAlaPheThrProAlaPheLeuMetGluPhe-----220
Db 235 GATTCAGTGCTTTCTCTCTGCTCCATGGTGGTGGATTACAAACACAGCAGTGGGTTTTC 294
Qy 221 -----LeuAsnSerTyrHis 225
Db 295 AATGATACTATTGATCCAGACAAAGATGGTTCGATTTGGACTTTGGACTTATAACTCTCTGAA 354
Qy 226 GlnGlnAlaAspTyrLeuSerThrHisTyrAlaGluGlnGlnAsnHisArgLeuPheGlu 245
Db 355 AGTGAAGACCAATACATTAGC-----TATTCTTGTGATGGTGTACTTTACTGAA 408
```



```
QY 246 AlaGlnArgAsnLeuPheAlaGlyValSerPheProGluPheLysAsp----- 261
D 249 TACCAAAAGAACCTGTTTGTAGTGCACCACTCCACTCAATTCAGAGATCCAAAGGTGTC 468
QY 262 -----SerProArgTrpArgGlnThrGlyIleSerValLeuAsnThrGluIle 277
D 269 TGGTATGAACCTTCTCAAAATATGGATTATGAGC-----GCTGCCAAATACCAAGACTAC 522
QY 278 LysLysGlnValTyAlaAsp----- 284
D 523 AAAATTTGAAATTTACTCCCTCTGATGACTTGAAGTCTGGAAGCTAGAATCTGCAATTGGCC 582
QY 285 -----GlyMetGlnPheGluLeu----- 290
D 583 AACGAAGGTTTCTTAGGCTACCAATACGAATGCCAGGTTCAGGTTTGATTGAAGTCCCACTGAG 642
QY 291 -----SerProIleTyrHisValAlaAlaIleAspIle----- 301
D 643 CAAGATCTTCCAAATCTTATTTGGGTTCATGTTTATTTCTATCAACCAGGTGCACCTGCT 702
QY 302 -----PheLeuLysAlaTyr 306
D 703 GCGGCTTCTCAACCAATATTTGTTGGATCCTTCAATGCTACTCATTTTGAAGCGTTT 762
QY 307 GlySerAlaLysArgValAsnLeuGluLysGluPheProGlnSerTyr-----ValGln 324
D 763 GACAATCAATCAGAGTG-----GTAGATTTTGGTAAGGACTACTATCCTTTGCAA 813
QY 325 ThrValGluAsnMetIleMetAlaLeuIleSerIleSerLeuProAspTyrAsnThrPro 344
D 814 ACTTCTTCAAC-----ACTGACCAACCTACCGTTTCAGCA 849
QY 345 MetPheGlyAspSerTrpIleThrAspLysAsnPheArgMetAlaGlnPheAla----- 362
D 850 TTA---GGTATTGCCTGGGCTTCA-----AACTGGAGTACAGTGCCTTTGTCCCAACT 900
QY 363 -----SerTrpAlaArgValPheProAlaAsnGlnAlaIleLys 375
D 901 AACCCATGGAGATCATCCATGCTTTTGGTCCGCAAGTTCCTTTTGAAC-----ACTGAA 954
QY 376 TyrPheAla-----ThrAspGlyLysGlnGlyLysAlaProAsnPheLeuSerLys 392
D 955 TATCAAGCTAATCCAGAGACTGAATGTATCAATTTGAAGCCGCAACCAATATTTGAAC--- 1011
QY 393 AlaLeuSerAsnAlaGlyPheTyrThrPheArgSerGlyTyrAspLysAsnAlaThrVal 412
D 1012 ---ATTAGTAATGCTGGT-----CCCTGGTCTCGTTTGTCTACTAAC 1050
QY 413 MetValLeuLysAlaSerProProGlyGluPheHisAlaGlnProAspAsnGlyThrPhe 432
D 1051 ACACTCTAATCAAGCCCAATCTTACAATGCGATTTCAGCACTCGACTCGTACCCCTA 1110
QY 433 GluLeuPheIleLysGlyArgAsnPheThrProAspAlaGlyValPheValTyrSerGly 452
D 1111 GAGTTGAGTTG-----GTTTACGCTGTTAACACC 1140
QY 453 AspGluAlaIleMetLys-----LeuArgAsnTrpTyrArgGlnThrArg 467
D 1141 ACACAAACCATATCCAAATCCGCTTTTGGCGACTTATCATCTTGGTTCAAGGGTTTAGAA 1200
QY 468 -----IleHisSerThrLeuThrLeuAsp 475
D 1201 GATCTGGAAGATATATTGGAATGGGTTTGAAGTCAGTCTTCTTCTCTTTTGGAC 1260
QY 476 AsnGlnAsnMetValIleThrLysAlaArgGlnAsnLysTrpGluThrGly----- 492
D 1261 CGTGGTAACTCAAGTCAAGTTTGTCAAGAGAACCCATATTTCAACAACAGAAATGTCT 1320
QY 493 -----AsnAsnLeuAspValLeuThrTyrThrAsnProSerTyrProAsn----- 507
D 1321 GTCAACAACCAACCAATTCAGTCTGAGAACGACCTAAGTTTACTATATAAGTGTACGCGCTA 1380
QY 508 LeuAspHisGlnArgSerValLeuPheIleAsn-----LysLys 520
```

```
Db 1381 CTGGATCAAAAACATCTTGGAAATTGACTTCAACGATGGAGATGTGTTCTTCAAATACC 1440
QY 521 TyrPheLeuValIleAspArgAlaIleGlyGluAlaThrGlyAsnLeuGlyVal 538
D 1441 TACTTCATGACCACCGGTAACGCTCTAGGATCTGTGAACATGACCACTGGTGTC 1494

RESULT 9
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 276 Length: 4403765
Score: 120.50 Matches: 85
Percent Similarity: 34.20% Conservatives: 60
Best Local Similarity: 20.05% Mismatches: 150
Query Match: 3.45% Indels: 129
DB: 3 Gaps: 22

US-09-802-285A-2 (1-659) x US-09-103-840A-2 (1-4403765)
QY 131 AspAsnGluValArgTrpGlnLeuHisArgValLysTrpTrpGlnAlaMetAla----- 148
D 3679590 GACAAGGCTATCAGTGG---CTGCACACCGTGGGCGCCGAGAAATGCCAAGCCGTGG 3679534
QY 149 LeuValTyrHisAlaThrGlyAspGluLysTyrAlaArgGluTrpValTyrGlnTyrSer 168
D 3679533 ATGCTGTACTACGACCCGCGC---GCCACCCAGCCGCCACACCATGATTCAG----- 3679483
QY 169 AspTrpAlaArgLysAsnProLeuGlyLeuSerGlnAspAsnAspLysPheValTrpArg 188
D 3679482 GAATGGCGCGACAAGTACCGAGGTGATGATGATGATGATGATGATGATGATGATGATGAT 3679423
QY 189 ProLeuGluValSerAspArgValGlnSerLeuProThrPheSerLeuPheValAsn 208
D 3679422 ACATTGCAACGCGCAAAAGCGACTCGGATCATTCACCCGCGCGCAACTC---ACCGAG 3679366
QY 209 SerProAlaPheThrProAlaPheLeuMetGluPheLeuAsnSerTyrHisGlnAla 228
D 3679365 CGGCCCGACCTATTCGCCGCTGG----- 3679342
QY 229 AspTyrLeuSerThrHisTyrAlaGluGlnGlyAsnHisArgLeuPheGluAlaGlnArg 248
D 3679341 GACAGTATGTCG-----GAGCGCGCAAAACCGCTCTTTGCCCGCCAGATG 3679297
QY 249 AsnLeuPheAlaGlyValSer----- 255
D 3679296 GAGGTGTTGCGCGGTTCTTCGAAAAATCGGACTCGAATGTTGGCCGCTGCTGACCGCG 3679237
QY 256 PheProGluPheLysAspSerProArg-----TyrArgGlnThrGly 269
```

Db	3679236	ATCGAGGATCTCGCGAGTCGGACAACACAGTGTGGTGTCTACATCTGTGGCGGCACATGGC	3679177
Qy	270	IleSerValLeuAsnThr-----GluIleLysLysGlnValTyrAlaAspGlyMet	286
Db	3679176	GCCAGCATGGAGGGCACCAACACCGTTCGTTCAATGAGATGAGTTCCTTAAACGGCTG	3679117
Qy	287	-----GlnPheGluLeuSerProIleTyr--HisValAlaAlaIle	299
Db	3679116	GATCTGGATCGCGAGCGCGCAATTGGAGCTGATCGAAACAATACGCGCGCATCGCGGCACATC	3679057
Qy	300	-----AspIlePheLeuLysAlaTyrGlySerAlaLysArgValAsn	313
Db	3679056	GGCGACAGTTCACCGCACCGCATTTCCGACGCGGTGGCGCGCATCGGACACACCCCG	3678997
Qy	314	LeuGlu-----LysGluPheProGlnSerTyrValGlnThrValGluAsnMetIleMet	331
Db	3678996	TTGCAGTGGGCAAGCAGATGGCCAGCCACCTGGCGCGCACGCGCATCATTTGGTGGTC	3678937
Qy	332	Ala-----	332
Db	3678936	GCITGGCGCGCCGAGTCGCGCCAGACGCGCGTTCGTYAGCCAGTTTCACCACTGCATC	3678877
Qy	333	-----LeuIleSerIleSerLeuProAsp-----	340
Db	3678876	GACATCGCGCGCAGCGTGTGGCGGCATCGGTTTACCGGAGCGGCCCATGTCTCGACGC	3678817
Qy	341	---TyrAsnThrProMetPheGlyAspSerTrpIleThrAspLysAsnPheArgMetAla	359
Db	3678816	TTCCAGCAGGAACGATGGACGGGAACAGTTTCGTG-----	3678781
Qy	360	GlnPheAlaSerTrpAlaArgValPheProAlaAsnGlnAla-----Ile	374
Db	3678780	-----CGACCTTCGACGCGTGAACCGCAAGACGCCACCGCATCCGGTG	3678739
Qy	375	LysTyrPheAlaThrAspGlyLysGlnGlyAlaProAsnPheLeuSerLysAlaLeu	394
Db	3678738	CAGTACTTC-----GAAAACTTCGCGACCGCTGCATC	3678706
Qy	395	SerAsnAlaGlyPheTyrThrPhe-----ArgSerGlyTrpAspLysAsnAla	410
Db	3678705	TACAAAGACGCGTGGTGGCGTGGCTGCTGGACGAAGCGCCCTGGGATCTGTCAACCG	3678646
Qy	411	ThrValMetValLeuLysAlaSerProProGlyGluPheHisAlaGlnProAspAsnGly	430
Db	3678645	GAGACGATG-----CGACGGTTCGCGCCGGGACCTAC-----GACCCGCGACCGGAC	3678598
Qy	431	ThrPheGlnLeuPheIleLysGlyArgAsnPheThrProAspAlaGlyValPheValTyr	450
Db	3678597	GTCTGGGAGCTGTACTACTTACCAGATGACTTCTCCAGCGCAAAAACCTTGGACGCGGAG	3678538
Qy	451	SerGlyAspGluAlaIleMetLysLeuArgAsnTrpTyrArgGlnThrArgIleHisSer	470
Db	3678537	CATCCGACAGGTGCGCGAGTCCACCCAGTGTGTGGCAGGAGGCCGCAACACCGG	3678478
Qy	471	ThrLeuThrLeu	474
Db	3678477	GTCTGCGCGCTG	3678466

## RESULT 10

```

US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FRASER, Owen R.
; APPLICANT: WHITE, Claire M.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2

```

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.:      276      Length:      4411529
Score:          120.50    Matches:      85
Percent Similarity: 34.20% Conservative: 60
Best Local Similarity: 20.05% Mismatches: 150
Query Match:      3.45% Indels:      129
DB:               3      Gaps:       22

US-09-802-285A-2 (1-659) x US-09-103-840A-1 (1-4411529)

QY      131 AspAsnGluValArgTrpGlnLeuHisArgValIysTrpTrpGlnAlaMetAla----- 148
      |||
      :|:
      :|:
Db      3684804 GACAAGGCTATCGAGTGG---CTGCACACCGTGGGGCCCGAAGATGCCACCAAGCCGTGG 3684748
      |||
      :|:
      :|:
QY      149 LeuValTyrHisAlaThrGlyAspGluTyrAlaArgGluTrpValTyrGlnTyrSer 168
      |||
      :|:
      :|:
Db      3684747 ATGCTGACTACGGCACCGGC---GCCACCCACGGCCACACACGCTATTTCAG----- 3684697

QY      169 AspTrpAlaArgIysAsnProLeuGlyLeuSerGlnAspAsnAspIysPheValTrpArg 188
      |||
      :|:
      :|:
Db      3684696 GAATGGCCGCACCAAGTACCGAGGTGAGTTCGATGATGGTGGGATGTGTACCGGCAGAAG 3684637
      |||
      :|:
      :|:
QY      189 ProLeuGluValSerAspArgValGlnSerLeuProThrPheSerLeuPheValAsn 208
      |||
      :|:
      :|:
Db      3684636 ACATTCAACCGGCAAAACGCACTCGGGATCATTTCCACCCGACCGCCGAATC---ACCGAG 3684560
      |||
      :|:
      :|:
QY      209 SerProAlaPheThrProAlaPheLeuMetGluPheLeuAsnSerTyrHisGlnGlnAla 228
      |||
      :|:
      :|:
Db      3684579 CGGCCCGACCTATTCCTCCGGCGTGG----- 3684556

QY      229 AspTyrLeuSerThrHisTyrAlaGluGlnGlyAsnHisArgLeuPheGluAlaGlnArg 248
      |||
      :|:
      :|:
Db      3684555 GACAGTATGTGG-----GAGCGCAAAACCGGTCTTTTGCCGCCAGATG 3684511
      |||
      :|:
      :|:

```

QY 341 --TyrAsnThrProMetPheGlyAspSerTrpIleThrAspLysAsnPheArgMetAla 359  
Db 3684030 TTCGACGAGAACCGATGACGAGAACCGATTTTCGTG----- 3683995  
QY 360 GlnPheAlaSerTrpAlaArgValPheProAlaAsnGlnAla-----Ile 374  
Db 3683994 -----CGGACCTTCGACGCGCTGACGCGGAGAACCGCCACACCGGTG 3683953  
QY 375 LysTyrPheAlaThrAspGlyLysGlnGlyLysAlaProAsnPheLeuSerLysAlaLeu 394  
Db 3683952 CAGTACTTC-----GAAACTTCGGCAGCGGTGCATC 3683920  
QY 395 SerAsnAlaGlyPheTyrThrPhe-----ArgSerGlyTrpAspLysAsnAla 410  
Db 3683919 TACAAAGACGGTGTGGGGCGTGGCTGCTGGACAGCGCCCTGGGATCTGTCAACG 3683860  
QY 411 ThrValMetValLeuLysAlaSerProProGlyGluPheHisAlaGlnProAspAsnGly 430  
Db 3683859 GAGACGATG-----CGACGGTTGCGCGCGGGACCTAC-----GACCCGACCAAGGAC 3683812  
QY 431 ThrPheGluLeuPheIleLysGlyArgAsnPheThrProAspAlaGlyValPheValTyr 450  
Db 3683811 GTCGGGAGCTGTACTACCTACGATGACTTCTCCAGCGGAAACCTGGCAGCGGAG 3683752  
QY 451 SerGlyAspGluAlaIleMetLysLeuArgAsnTrpTyrArgGlnThrArgIleHisSer 470  
Db 3683751 CATCCCGACAGGTGCGCGGCTCACCCAGCTGTGGTGGCAGGCGCGAAACGACCGG 3683692  
QY 471 ThrLeuThrLeu 474  
Db 3683691 GTGCTGCGCGCTG 3683680

## RESULT 11

US-09-543-681A-3009  
; Sequence 3009, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; FILE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709,1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 3009  
; LENGTH: 3087  
; TYPE: DNA  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-3009

## Alignment Scores:

Pred. No.:	0.00468	Length:	3087
Score:	117.50	Matches:	167
Percent Similarity:	32.80%	Conservative:	101
Best local Similarity:	20.44%	Mismatches:	280
Query Match:	3.36%	Indels:	269
DB:	4	Gaps:	49

US-09-802-285A-2 (1-659) x US-09-543-681A-3009 (1-3087)

QY 5 IlePheLysArgIleLeValPheAlaValIleAlaLeuSerSerGlyAsnIleLeuAla 24  
Db 817 ATCGACAAATAATTAGCATTTAGGAAAAAATTAGCCTTAGAACAA-----CAT 864  
QY 25 GlnSerSerIleThrArgLysAspPheAspHisIleAsn-----LeuGlu 40  
Db 865 GCTGATGGCTCAATAACCCCAAGCCCTTGATCATCTTAATCGCAGAAATTTTATCAA 924  
QY 41 TyrSerGlyLeu-----GluLysValAsnLysAlaValAlaAlaGlyAsn---TyrAsp 57

Db 925 GTCGAGGGGGTCTTTAGTGAAGAAACCCAAAAAGCATTTACTTGATGCCAACATGCTAAGA 984  
QY 58 AspAlaAlaIleAlaLeuLeu-----AlaTyrTyrArgGluLysSerLysAlaArg 74  
Db 985 GATGTAGGCAAAACGCTTCTTCAACAGCGATTTACTTGGTAGTCGCTTCCTTATCA--- 1041  
QY 75 GluProAspPheSerAsnAlaGluLysProAlaAspIleArgGlnProIleAspLysVal 94  
Db 1042 -----GCAATCGATAGAAATAAATTAGAACGCTC 1071  
QY 95 -----ThrArgGluMetAlaAspLysAlaLeuVal-----HisGln 106  
Db 1072 TATTTATTTAGGCACCTCGTTACGTTCTTGAACAAGGTTTCACACGGGGAGTGGTATCCAA 1131  
QY 107 PheGlnProHisLysGlyTyr-----GlyTyrPheAsp-----TyrGlyLys 120  
Db 1132 ATTATTTACTCAGTAGTTATCAACACAGAGAGCTTTTGTATGCATGGTTTATTGGTCGC 1191  
QY 121 AsnIle-----AsnTrpGlnMetTrpPro--- 128  
Db 1192 CATATTTAGCAAAACATAATCTGTTAGCACCGACACAAACAGCCATGATGTTGTTAAAC 1251  
QY 129 -----ValLysAspAsnGluVal----- 134  
Db 1252 GCAACAGCGCGCATATTTGAAAAGATAATGAATCGTTGATGCCAAATGTCGATATTCTC 1311  
QY 135 ArgTrpGlnLeuHisArgValLysTrpTrpGlnAlaMetAlaLeuValTyrHisAlaThr 154  
Db 1312 AATACCAATTACAG-----TGGATGATAAAAGTTTATTG-----ATGTTG 1353  
QY 155 GlyAspGluLysTyrAlaArgGluTrpValTyrGlnTyrSerAspTrpAlaArgLysAsn 174  
Db 1354 CCAGATTATCAACAGCGCCACCAAGCCTTAGCACAAATTGCAAAAGTTGGTCAACAAACC 1413  
QY 175 ProLeu-----GlyLeuSerGlnAspAsnAspLysPhe----- 185  
Db 1414 ATCTTAGTTCTAAAGGTGTCGCTGGCTTTAAATCTGATGGCTCTATTTTTCACCAT 1473  
QY 186 ---ValTrpArgProLeuGluValSerAspArgValGlnSerLeuProProThrPheSer 204  
Db 1474 TCACAACTATTATCCAGCTTATCCAAAGATGCATTTGGTGGCTTAGCGCTAGTGTAT 1533  
QY 205 LeuPheValAsnSerPro-----AlaPheThrProAla----- 215  
Db 1534 GCCTTAAGTCATTCCTCCCTTTTCGCTTTCATCACCAGCACATCGACGCTTAAAGATGTA 1593  
QY 216 PheLeuMetGluPheLeuAsnSerTyrHisGlnGlnAlaAspTyrLeuSerThrHisTyr 235  
Db 1594 TTATTAATAATCGGTATCTATACCAAGAAACACAAATTCCTCTGTTATTAAGTGGCAGA 1653  
QY 236 AlaGluGlnGlyAsnHisArgLeuPheGluAlaGlnArgAsnLeuPheAlaGlyValSer 255  
Db 1654 CATCTACGGGGTTACATAAA-----ATAACT 1680  
QY 256 PheProGluPheLysAspSerProArgTrpArgGlnThrGlyIleSerValLeuAsnThr 275  
Db 1681 ATTGATCCCTTCAAA-----TGG-----ATGCTCTTCAGGTAAG 1716  
QY 276 GluIleLysLysGlnValTyrAlaAspGlyMetGlnPheGluLeuSerProIleTyrHis 295  
Db 1717 -----CCCGATGGTAAACAA---GAGCTAGATACTACA----- 1746  
QY 296 ValAlaAlaIleAspIlePheLeuLysAlaTyrGlySerAlaLysArgValAsnLeuGlu 315  
Db 1747 TTAGCAGCGCTTATGCAAAATTAGCAACAAAGATAGCTTTGAAGGTATTAAACGACAGAA 1806  
QY 316 LysGluPheProGlnSerTyrValGlnThrValGluAsnMet---IleMetAlaLeuIle 334  
Db 1807 AATGACCGGTAGGTGCATGGCAATGAATTTATGTTCAATGGCAATCCAGCGTAGACGA 1866  
QY 335 SerIleSerLeuProAspTyrAsnThrProMetPheGlyAspSerTrpIleThrAspLys 354  
Db 1867 TCATTAACAGGCCACAG-----CAAGCTGG----- 1893

```
Qy 355 AsnPheArgMetAlaGlnPheAlaSerTrpAlaArgValPheProAlaAsnGlnAlaIle 374
Db 1894 -----CTTGCTATCGTCGTCGGTGTATTAGTCGTTATTAGTCGGTAATGAAAGT--- 1941
Qy 375 LysTyrPheAlaThrAspGlyLysGlnGlyLysAlaProAsnPhe----- 389
Db 1942 -----TATGAGATAACAACCGTTATGGTCGTTACTTGCNATATGGCAACTTGAATT 1995
Qy 390 LeuSerLysAlaLeuSerAsnAlaGlyPheTyrThrPheArgSerGlyTrpAspLysAsn 409
Db 1996 ATTCCTCGGATTTAACTAAATCTCGTTTT-----AGTCATGCAGGTGGGATTGGAAT 2049
Qy 410 -----AlaThrValMetValLeuLysAlaSerProProGlyGluPheHisAla 425
Db 2050 CGATATCCAGGACGACTACGATCCACCTT-----CCTATGATGAACTAGAAGCA 2100
Qy 426 Gln-----ProAspAsnGlyThrPheGluLeuPheLeuLysGlyArgAsnPhe 441
Db 2101 AAATAAGCCNATGCCAAGTCAGGCATAGAGAAGATGTTGCTTCAACACACAGT--- 2157
Qy 442 ThrProAspAlaGlyValPheValTyrSerGlyAspGluAlaIle----- 456
Db 2158 -----TATTCGTGTCCTCAATACACATAACATAATAATAACAGTATG 2193
Qy 457 -----MetLysLeu-----ArgAsnTrpTyrArgGlnThrArgIleHis----- 469
Db 2194 TTTCCCATGAATTAACATGTCACAGTAATATCAACAACAAAGTCTAAGAGCGAATAAA 2253
Qy 470 SerThrLeuThrLeuAspAsn-----TATTCGTGTCCTCAATACACATAACAGTATG 2193
Db 2254 TCCTATTCTTATTGATAATAGATGATCGCATAGCTGCTGCTGTTGAAATAACGAT 2313
Qy 477 -----GlnAsnMet 479
Db 2314 AAACAACATACAACAGAAACGACACTGTTCAGATTGCTGTCCTAAGTTACATCAATT 2373
Qy 480 ValIleThrLysAlaArgGlnAsnLysTrpGluThrGly-----AsnAsnLeuAsp 496
Db 2374 ATAATTAATGTAAGAAAGTGAACCAACTCGGTACTCAATTAACCTTAATATGATGCGAT 2433
Qy 497 ValLeuThrTyrThrAsnProSerTyrProAsnLeuAspHisGlnArgSerValLeuPhe 516
Db 2434 ACATTAATTGATCCGCGGGTAACTTGTATAAATTAGCTAAGGCGCAACGGTAGAATT 2493
Qy 517 IleAsnLysLysTyrPheLeuValIleAspArg----- 527
Db 2494 AGTTATCAAAAACAATACTCTGTGTATGATGACAGAAATTCACAACAACAGACAATGTTT 2553
Qy 528 -----AlaIleGlyGluAlaThrGlyAsnLeuGlyValHisTrpGlnLeu 542
Db 2554 GCAACACCGTTATCTCTCAGGTAAAGCCCTAAAATGCAAAATTAATATGATATGCAATA 2613
Qy 543 LysGluAspSerAsnProValPheAspLysThrLysAsnArgValTyrThrTyrArg 562
Db 2614 GCCATAGAAGCA-----CAAGATAATAAAGCGCTGAATACACCGTATTGCGAG 2661
Qy 563 AspGlyAsnAsnLeuMetIleGlnSerLeuAsnAlaAspArgThrSerLeuAsnGlu 582
Db 2662 CATATAAT-----CACTTCATGCTAAGATAAAGATAAATAACCGCAAGAA 2706
Qy 583 GluGlyLysValSerTyrValTyrAsnLysGluLeuLys----- 595
Db 2707 GAGGGA-----TATGCTTTTTTAAATGCCACCGAAGCAATTCATCTCAAGCTTTA 2757
Qy 596 -----ArgProAlaPheValPheGluLysProLysLysAsnAlaGlyThr 610
Db 2758 TTATTATCAAGTATGATCCCACTATGATGTAATGTTAAATAAATAAATAAATAAATAA 2808
Qy 611 GlnAsnPheValSerIleValTyrPro-----TyrAspGlyGlnLysAlaPro 626
Db 2809 CAATTAACGCTAAGTATTGTTAACTGATTTAACTTAATTAATCAAGGTATTGAAGCTGAC 2868
```

627 GluIleSerIleArgGluAsnLysGlyAsnAspPheGluLysGly----- 641

2869 CAATTT-----GATAATTAAGGTAAACCAAGTTGAAGTGTATTATTCTCGCA 2919

642 -----LysLeuAsnLeuThrLeuThrIleAsnGly 651

2920 TGGCTTACGCCGATCTCAACCAATAAGTAGTACAGTTACTTACTGTAAGGG 2970

RESULT 12

US-08-956-171E-268/c

; Sequence 268, Application US/08956171E

; Patent No. 6593114

GENERAL INFORMATION:

APPLICANT: Charles Kunsch

Gil H. Choi

Patrick S. Dillon

Craig A. Rosen

Steven C. Barash

Michael R. Fannon

TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences

NUMBER OF SEQUENCES: 5256

CORRESPONDENCE ADDRESS:

ADDRESSEE: Human Genome Sciences, Inc.

STREET: 9410 Key West Avenue

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/956,171E

FILING DATE: 20-Oct-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/009,861

FILING DATE: January 5, 1996

APPLICATION NUMBER: 08/781,986

FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:

NAME: Mark J. Hyman

REGISTRATION NUMBER: 46,789

REFERENCE/DOCKET NUMBER: PB248P1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (240) 314-1224

TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 268:

SEQUENCE CHARACTERISTICS:

LENGTH: 4702 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 268:

US-08-956-171E-268

Alignment Scores:

Pred. No.: 0.0138 Length: 4702

Score: 116.00 Matches: 138

Percent Similarity: 33.20% Conservative: 103

Best Local Similarity: 19.01% Mismatches: 301

Query Match: 3.32% Indels: 185

DB: 4 Gaps: 27

US-09-802-285A-2 (1-659) x US-08-956-171E-268 (1-4702)

Qy 4 LysIlePheLysArgIleValPheAlaValIleAlaLeuSerSerGlyAsnIleLeu 23

Db 3000 AAAATATTCCAGGAGTTTCTAGTTATGCAATGATTAAGATGAGTCCAGAGCAATCAGA 2941

Qy 24 AlaGlnSerSerIleThrArgLysAspPheAspHisIleAsnLeuGluTyrSerGly 43

Db 2940 GCAAAATCGCAATCTTACGGCAAGGTTCA---GACCAAAATCCGTCAAATTTATCTCAT 2884  
Qy 44 LeuGluLysValAsnLysAlaValAlaAlaGlyAsnTyrAspAlaAlaLysAlaLeu 63  
Db 2883 TTAACMCGTGCACAGGTGAATTCAGCG---AACTGGGAAGTCAAGCT----- 2836  
Qy 64 LeuAlaTyrArgGluLysSerLysAlaArgGluProAspPheSerAsnAlaGluLys 83  
Db 2835 TTCAGCCGTTTCGAGAGCAATTCACAACTTAGTCCT-----AAAGTAGAAAAA 2785  
Qy 84 ProAlaAspIleArgGlnProIleAspLysValThrArgGluMetAlaAspLysAlaLeu 103  
Db 2784 TTTGCACAAATTATTAGAGAAATTAACACAAATGAATAGACACTGCTGACCCGTTCAA 2725  
Qy 104 ValHisGlnPheGlnProHisLysGlyTyrGlyTyrPheAspTyrGlyLysAspIleAsn 123  
Db 2724 GAACAGACCAACAACTTCTTAATAATTCGGT----- 2692  
Qy 124 TrpGlnMetTrpProValLysAspAsnGluValArgTrpGlnLeuHisArgValLysTrp 143  
Db 2691 -----TTGCAATAAGCATTTCTGAAATTCGCAAAAGTCACAT-TTTCTAATG 2648  
Qy 144 TrpGlnAlaMetAlaLeuValTyrHisAlaThrGlyAspGlu-----LysTyrAlaArg 161  
Db 2647 TGGCTTTCCTATCATCTTTTTTAAGAAAACAACACTGAAAGGAATAAGCATGAAAGAAA 2588  
Qy 162 GluTrpValTyr----- 165  
Db 2587 AATTGGATTATGCAATTAATGTCACCTTTAATTATTATAATTGCCATAGTTAGTATGATA 2528  
Qy 166 -----GlnTyrSerAspTrpAlaArgLysAsnProLeuGlyLeuSerGln 180  
Db 2527 TTTTGTTCACAAATAATGAGATCAATCAGAAAAGGATCCCAAAGTGTAAGTAAT 2468  
Qy 181 AspAsnAspLysPheValTyrArgProLeuGluValSerAspArgValGlnSerLeuPro 200  
Db 2467 AAAATAATAA-----ATACATATCGCAATTTGTTAACGAGATCAACCA 2423  
Qy 201 ProThrPheSerLeuPheValAsnSerProAlaPheThrProAlaPheLeuMetGluPhe 220  
Db 2422 ACGCATATAAC-----GGTAAAAAGTTGAGCTGGTCAAGCATTTATTAAAAAGTTA 2369  
Qy 221 LeuAsnSerTyrHisGlnAlaAspTyrLeuSerThrHisTyrAlaGlu----- 237  
Db 2368 GCAAAATGAGAAAACTATAAATTTGAAAACAGTAAACAGAAACGTTGCTGAGTCTGGTTG 2309  
Qy 238 ---GlnGlyAsnHisArgLeuPheGluAlaGlnArgAsnLeuPheAlaGlyValSerPhe 256  
Db 2308 AAAAATGGTGGATACCAAGTCATGATTGTTATCCAGAAAACCTTTTCAAAATGGCAATG 2249  
Qy 257 ProGluPheLysAspSerProArg-----TrpArgGlnThrGlyIleSer--- 271  
Db 2248 CAATTAGACGCTAAACACCACTACGATAATCGCTACAGTATAAAAACAGCTGTAGACAA 2189  
Qy 272 -----ValLeuAsnThrGluIleLysLysGlnValTyrAlaAspGlyMetGln 287  
Db 2188 AAAGAAGAAGTAGCTAAACACACAGAAAAGTTGTAAGTAATGTACTTAACGACTTTAAC 2129  
Qy 288 PheGluLeuSerProIleTyrHisValAlaAlaIleAsp----- 300  
Db 2128 AAAAATCTAGTCAAAATTTATTACAGCATCATTTGATAATTTACATAATGCCAAAAA 2069  
Qy 301 -----IlePheLeuLysAlaTyrGly-----SerAlaLysArg 311  
Db 2068 AATGTTGGCGCTATTATGACGGGTGAACATGGTGTGAATAGTAATAATCTCGAATTACTTA 2009  
Qy 312 ValAsnLeuGluLysGluPheProGlnSerTyrValGlnThrValGluAsnMetIleMet 331  
Db 2008 TTAATTCOAATTAACGACTTCCCGAAATTTATTACAGATACCTGTTGAATTCATTTCT 1949  
Qy 332 AlaLeuIleSerIleSerLeuProAspTyrAsnThrProMetPheGlyAspSerTrpIle 351

Db 1948 GCACAAAGACATTACA----- 1931  
Qy 352 ThrAspLysAsnPheArgMetAlaGlnPheAlaSerTrpAlaArgValPheProAlaAsn 371  
Db 1931 ----- 1931  
Qy 372 GlnAlaIleLysTyrPheAlaThrAspGlyLysGlnGlyLysAlaProAsnPheLeuSer 391  
Db 1930 -----AAATGGTTCCAAACA-----TACAAAT 1910  
Qy 392 LysAlaLeuSerAsnAlaGlyPheTyrThrPheArgSerGlyTrpAspLysAsnAlaThr 411  
Db 1909 AAATCATTTTTCAGTGGCAATTCAGATACGTTTCAGAGTGAACACAGATATATATGTTTCG 1850  
Qy 412 ValMetValLysLysAlaSerProProGlyGluPheHisAlaGlnProAspAsnGlyThr 431  
Db 1849 ACTTTAATTGAAGAAA-----CAAAATTCATTAA 1823  
Qy 432 PheGluLeuPheIleLysGlyArgAsnPheThrProAspAlaGlyValPheValTyrSer 451  
Db 1822 TTTGAC-----GAGCACAATACACGATGCGATAAAATGTTTACAGATTATATAA 1775  
Qy 452 GlyAspGluAlaIleMetLysLeuArgAsnTrpTyrArgGlnThrArgIleHisSerThr 471  
Db 1774 TCGCAAAAAGATAGCTGGAACTTGATACTAT-----ATCAATGCATTA 1730  
Qy 472 LeuThrLeuAspAsnGln-----AsnMetValIleThrLysAlaArgGln 486  
Db 1729 AAACAGATGGACAGCAAAATGATCAACAATCAAGTATGATCAAGATACAGGTAAGAAGAA 1670  
Qy 487 AsnLysTyrGluThrGlyAsnAsnLeuAspValLeu-----ThrTyrThr 501  
Db 1669 TATAAACAACACTGTTAAAGAAAACCTTAGATAAAATAGAGAAATCATTCATCAACAGAG 1610  
Qy 502 AsnPro---SerTyrProAsnLeuAspHisGlnArgSerValLeuPheIleAsnLysLys 520  
Db 1609 TCACCATTTTCAAAAGGTATGATTGAAGANTATCGTAAGCAATTAACAGATCACTGCAA 1550  
Qy 521 TyrPheLeuValIleAspArgAlaIleGlyGluAlaThrGlyAsnLeu----- 536  
Db 1549 GATGAGCTTGCAATAAACAAAGACTTACAAGATGCGCTAAATAGCATTAATAATGAACAAT 1490  
Qy 537 -----GlyValHisTrpGlnLeuLysGluAspSerAsnProValPheAsp 551  
Db 1489 GCTCAATTCGCTGAAAACCTTAGAGAAAACACTTCATGATGATATGTCAAAGAACTGAT 1430  
Qy 552 LysThrLysAsnArgValTyrThrTyrArgAspGlyAsnAsnLeuMetIleGlnSer 571  
Db 1429 ACAGATACAACATTTATCTAT-----AACATGCTCAACAAGAC 1391  
Qy 572 LeuAsnAlaAspArgThrSerLeuAsnGluGluGlyLysValSerTyrValTyrAsn 591  
Db 1390 TTTTATAGCT-----GCAGGTTTAAATGAGGTGAAGCTAATAATACGAACTATTGTC 1337  
Qy 592 LysGluLeuLysArgProAlaPheValPheGluLysProLysLysAsnAlaGlyThrGln 611  
Db 1336 AAAGAGCAAAACGTTATAAAAATGAATATATTTGAAAAAACCGTTAGCAGAACACATT 1277  
Qy 612 AsnPheValSerIleValTyrProTyrAspGlyGlnLysAlaProGluIleSerIleArg 631  
Db 1276 AATTTTAACAGAT-----TACGATAACCAAGTTGCGAAGACACAAAGTAGTTTG 1229  
Qy 632 GluAsnLysGlyAsnAspPheGluLysGlyLysLeuAsnLeuThrLeuThrIleAsnGly 651  
Db 1228 ATTAATGATGTTGTCAAAAGTCACAGCTACTGAA-----ACGATTAAAAAGTAATGAT 1178  
Qy 652 LysGlnGlnLeuValLeu 657  
Db 1177 ATTAATCAATTAACGTGT 1160

RESULT 13

US-09-220-132-20

; Sequence 20, Application US/09220132



QY 603 LysPro-----LysLysAsnAlaGlyThrGlnAsn----- 612  
DB 10837 TTCCTGACCTTGGCCAGGAAGTGGCCCTGAATGCTAAACACTAAGAACACGAGATCAGA 10896  
QY 613 PheValSerIleValTyrProTyrAspGlyGlnLysAlaProGluIleSerIleArgGlu 632  
DB 10897 TGGAAATAGAGTCCGGATTCAATCTGGGTCTTCCAGAGCCAGGTCAGCTT----- 10950  
QY 633 AsnLysGlyAsnAspPheGluLysGlyLysLeuAsnLeuThrLeuThrIleAsnGly 651  
DB 10951 -----TCCAATGACCAAGAAAAGGCACACCTTGACATTCGAGGATCCTTAGAAGGA 11001  
RESULT 14  
US-07-973-320-3  
; Sequence 3, Application US/07973320  
; Patent No. 5286486  
; GENERAL INFORMATION:  
; APPLICANT: Payne, Jewel M.  
; APPLICANT: Fu, Jenny M.  
; TITLE OF INVENTION: No. 5286486el Bacillus thuringiensis Gene  
; TITLE OF INVENTION: Encoding a Coleopteran-Active Toxin  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: David R. Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: USA  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/973,320  
; FILING DATE: 19921106  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/788,638  
; FILING DATE: 6-NOV-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Saliwanchik, David R.  
; REGISTRATION NUMBER: 31,794  
; REFERENCE/DOCKET NUMBER: MA68.C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 904-375-8100  
; TELEFAX: 904-372-5800  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3414 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Bacillus thuringiensis  
; STRAIN: kumamotoensis  
; INDIVIDUAL ISOLATE: HD867  
; IMMEDIATE SOURCE:  
; LIBRARY: Landagem (TM)-11 library of J.M. Fu  
; CLONE: 867  
US-07-973-320-3  
Alignment Scores:  
Pred. No.: 0.00942 Length: 3414  
Score: 115.50 Matches: 135  
Percent Similarity: 35.99% Conservative: 100  
Best local Similarity: 20.67% Mismatches: 204  
Query Match: 3.31% Indels: 216  
DB: 1 Gaps: 37

US-09-802-285A-2 (1-659) x US-07-973-320-3 (1-3414)  
QY 48 AsnLysAlaValAla-----AlaGlyAsnTyrAspAspAlaAlaLysAlaLeu 63  
DB 361 AATAAAGCACTTGCAGATTAGAGGATTAGGAATAAC----- 399  
QY 64 LeuAlaTyrTyrArgGluLysSerLysAlaArgGluProAspPheSerAsnAlaGluLys 83  
DB 400 TTAACGATATATCAACAG-----GCACCTGAAGATTGGCTGAACAATCCTGATGAT 450  
QY 84 ProAlaAspIleArgGlnProIleAspLysValThrArgGluMetAlaAspLysAlaLeu 103  
DB 451 CCAGCAACTATAAACACGAGTAGATCGTTT-----CGTATATTAGAT---GCTTTTA 501  
QY 104 ValHisGlnPheGlnProHisLysGlyTyrGlyTyrPheAsp----- 117  
DB 502 TTTGAATCATATATGCCGTTCATTTAGGGTTGCTGATATGAATAATACCATTACTAACAGTT 561  
QY 118 TyrGlyLysAspIleAsnTyrGlnMetTyrProValLysAspAsnGluVal----- 134  
DB 562 TACGCACAAAGCGGCAACCTTCATCTAGCTTTATTAAAGAGATTCTACTCTTTATGAGAT 621  
QY 135 ArgTyrGlnLeu----- 138  
DB 622 AAATGGGAATTCACCTCAGAACAACTGAGGAAAATTATAATCGTCAAAAGAAACATATT 681  
QY 139 -----HisArgValLysTyrTyrGlnAla---MetAlaLeuValTyrHis 152  
DB 682 TCTGAATATTCCTAACCATTCGTTAAGTGGTATATAGTGGTCTTAGCAGATTGAACGGT 741  
QY 153 AlaThrGlyAspGlu-----LysTyrAlaArgGluTyrValTyrGlnTyr 167  
DB 742 TCCACTTATGAACAATGATGATAAATTATAATCGTTTCGTAGAGAAATGATATTATGTA 801  
QY 168 SerAspTyrAlaArgLysAsnProLeu-----GlyLeu 178  
DB 802 TTAGATATTGCTGCTGTTATTCCTATTATGACCTCGAATGATTCATGAGAAACAGT 861  
QY 179 SerGlnAspAsnAspLysPheValTyrArg-ProLeuGluValSerAspArgValGlnSe 198  
DB 862 AGCGAGTT-AACGAGAGAGTGTATACCGATCCAATTAGCTTGTCAATTAGCAATCCAGG 920  
QY 198 rLeuProThrPheSerLeuPheValAsnSerProAlaPheThrProAlaPheLeuMe 218  
DB 921 TATAGTCCAAAGTTTCTCAGATGGAATAACTCGAATTAGACACACACCTTGT--- 978  
QY 218 tGluPheLeuAsnSerTyrHisGlnGlnAlaAspTyrLeuSerThrHisTyrAlaGluGl 238  
DB 979 -----GATTATTAGATGAGCTTTATATATATAC 1007  
QY 238 nGlyAsnHisArgLeuPhe-----GluAlaGlnArgAsnLeuPhe-----Al 252  
DB 1008 ATCAAAATATAAGCATTTTCACATGAGATTCAACAGACCTATTATTATTGGAGTCACA 1067  
QY 252 aGlyValSerPheProGluPheLysAspSerProArgTyrPargGlnThrGlyIleSerVa 272  
DB 1068 TAAGGTTAGCTTTAAACAATCGAGCAATCCAAATTATAT---ACAACAGGCATATA--- 1119  
QY 272 lLeuAsnThrGluIleLysLysGlnValTyrAlaAspGlyMetGlnPheLeuSerPr 292  
DB 1120 -----TATCGTAAACAAGTGGATATATTTTCATCAGGGGCATATTCATTAGAGGTAATGA 1175  
QY 292 oIleTyrHisValAlaAla-----IleAspIlePheLeuLysAlaTyrGl 307  
DB 1176 TATCTATAGAAACATTAGCAGCTCCATCATGTTGTAGTTTATCCGTATACTCAGAAATTATGG 1235  
QY 307 Y-SerAlaLys----- 310  
DB 1236 TGTGAGCAAGTTGAGTTTTTACGTTGTAAGGCGACGTCATTATAGAGAGAGATAACAA 1295  
QY 311 -----ArgValAsnLeuGluLysGluPheProGlnSerTyrValGlnThrValG 327







This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 27, 2004, 17:26:51 ; Search time 49 Seconds  
(without alignments)  
4211.605 Million cell updates/sec

Title: US-09-802-285A-2

Perfect score: 3494  
Sequence: 1 MTKIKFRIIVFAVIALSSG.....KGKINLTITNGKQQLVLP 659

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1288442 seqs, 313154207 residues

Total number of hits satisfying chosen parameters: 1288442

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

```
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	243	7.0	896	12 US-10-282-122A-45584	Sequence 45584, A
2	130.5	3.7	681	12 US-10-282-122A-71561	Sequence 71561, A
3	128	3.7	25	9 US-09-802-285-3	Sequence 3, Appli
4	128	3.7	25	14 US-10-231-337-3	Sequence 3, Appli
5	126.5	3.6	696	12 US-10-282-122A-70746	Sequence 70746, A
6	123	3.5	532	15 US-10-369-493-22074	Sequence 22074, A
7	120	3.4	677	15 US-10-369-493-23212	Sequence 23212, A
8	118.5	3.4	579	15 US-10-369-493-23215	Sequence 23215, A
9	118.5	3.4	579	15 US-10-369-493-23221	Sequence 23221, A
10	117.5	3.4	1484	12 US-10-335-977-8966	Sequence 8966, A
11	117.5	3.4	2902	12 US-10-282-122A-59042	Sequence 59042, A
12	117	3.3	1163	15 US-10-452-024-107	Sequence 107, App
13	117	3.3	1250	10 US-09-769-736-10	Sequence 10, Appl
14	116	3.3	919	12 US-10-282-122A-70737	Sequence 70737, A
15	116	3.3	4563	10 US-09-802-640-32	Sequence 32, Appl

16	116	3.3	4563	12 US-10-403-902A-32	Sequence 32, Appl
17	115.5	3.3	987	12 US-10-282-122A-48445	Sequence 48445, A
18	115.5	3.3	1314	15 US-10-369-493-1532	Sequence 1532, Ap
19	115.5	3.3	1331	9 US-09-801-368-370	Sequence 370, App
20	114.5	3.3	645	14 US-10-130-973A-8	Sequence 8, Appli
21	114.5	3.3	1279	10 US-09-882-227-388	Sequence 388, App
22	114	3.3	871	15 US-10-369-493-13471	Sequence 13471, A
23	114	3.3	899	15 US-10-369-493-5864	Sequence 5864, Ap
24	114	3.3	2234	12 US-10-282-122A-46565	Sequence 46565, A
25	113.5	3.2	1398	12 US-09-841-553-8	Sequence 8, Appli
26	113.5	3.2	1398	13 US-10-090-624-6	Sequence 6, Appli
27	113	3.2	2179	14 US-10-224-999A-3481	Sequence 3481, Ap
28	111.5	3.2	846	15 US-10-320-797-3302	Sequence 3302, Ap
29	111.5	3.2	962	12 US-10-282-122A-77946	Sequence 77946, A
30	111	3.2	723	15 US-10-369-493-10942	Sequence 10942, A
31	110.5	3.2	563	9 US-09-815-242-5665	Sequence 5665, Ap
32	110.5	3.2	578	9 US-09-815-242-12206	Sequence 12206, A
33	110	3.1	4563	9 US-09-870-759-128	Sequence 128, App
34	110	3.1	4563	10 US-09-751-708A-128	Sequence 128, App
35	109.5	3.1	966	15 US-10-099-322-72	Sequence 72, Appl
36	109.5	3.1	966	15 US-10-044-564-72	Sequence 72, Appl
37	109.5	3.1	1465	12 US-10-282-122A-47271	Sequence 47271, A
38	109	3.1	634	12 US-10-282-122A-52070	Sequence 22070, A
39	107.5	3.1	1336	10 US-09-934-455-22	Sequence 22, Appl
40	107.5	3.1	1336	12 US-10-412-699B-654	Sequence 654, App
41	107.5	3.1	1336	12 US-10-225-066A-576	Sequence 576, App
42	107.5	3.1	1336	14 US-10-278-173-128	Sequence 128, App
43	107.5	3.1	1336	15 US-10-225-067-92	Sequence 92, Appl
44	107.5	3.1	1336	15 US-10-374-780A-204	Sequence 204, App
45	107	3.1	477	12 US-10-335-977-9400	Sequence 9400, Ap

#### ALIGNMENTS

#### RESULT 1

```
US-10-282-122A-45584
; Sequence 45584, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
```



QY 457 MRLNRVYQTHSTLTLDNQNMVITKARQNKWETGNNLDVLTYPNPSYPLNDHORSVLF 516  
Db 480 QKLAKGLNQVGLGKVTGIDLPNETIGQIEPLTNNSGNVLDLSIGQDYTSPIQLSQYVST 539  
QY 517 INKXVFLV--IDRAIGATGNLGVHWQLKEDSNPFVDFKTKNRVYTTTYRDGNNLMIQSLN 573  
Db 540 IANDGYRIQPHIGLAHDATNS-----DDIGPVKOKIKGNVLNKNVNSEDEIKVQK 591  
QY 574 ADRTSLNEEG 584  
Db 592 GFEMAFNEKG 602

RESULT 3  
US-10-282-285-3  
; Sequence 3, Application US/09802285  
; Patent No. US20020122793A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Dongfang  
; APPLICANT: Pojasek, Kevin  
; APPLICANT: Shriver, Zachary  
; APPLICANT: Holley, Kristine  
; APPLICANT: El-Shabrawi, Yosuf  
; APPLICANT: Venkataraman, Ganesh  
; APPLICANT: Sasisekharan, Ram  
; TITLE OF INVENTION: Heparinase III and Uses Thereof  
; FILE REFERENCE: M0656/7063HCL  
; CURRENT APPLICATION NUMBER: US/09/802,285  
; CURRENT FILING DATE: 2001-03-08  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Flavobacterium heparinum  
US-09-802-285-3

Query Match 3.7%; Score 128; DB 9; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 QVYADGMQFELSPIYHVAADIFLK 304  
Db 1 QVYADGMQFELSPIYHVAADIFLK 25

RESULT 4  
US-10-291-337-3  
; Sequence 3, Application US/10291337  
; Publication No. US20030099628A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Dongfang  
; APPLICANT: Pojasek, Kevin  
; APPLICANT: Shriver, Zachary  
; APPLICANT: Holley, Kristine  
; APPLICANT: El-Shabrawi, Yosuf  
; APPLICANT: Venkataraman, Ganesh  
; APPLICANT: Sasisekharan, Ram  
; TITLE OF INVENTION: Heparinase III and Uses Thereof  
; FILE REFERENCE: M0656/7063HCL  
; CURRENT APPLICATION NUMBER: US/10/291,337  
; CURRENT FILING DATE: 2002-11-08  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Flavobacterium heparinum  
US-10-291-337-3

Query Match 3.7%; Score 128; DB 14; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.00016;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 280 QVYADGMQFELSPIYHVAADIFLK 304  
Db 1 QVYADGMQFELSPIYHVAADIFLK 25

RESULT 5  
US-10-282-122A-70746  
; Sequence 70746, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 70746  
; LENGTH: 696  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-10-282-122A-70746

Query Match 3.6%; Score 126.5; DB 12; Length 696;  
Best Local Similarity 18.6%; Pred. No. 0.039;  
Matches 139; Conservative 109; Mismatches 284; Indels 217; Gaps 31;  
QY 9 IIVFAVIALSSGNI-LAQSSSITR--KDFDHINLEYSGLEKVNKAVAAGN-----Y 56  
Db 30 VFIFAVVLRLGYLOIAQSGSHYKQLIKNDENIT-----VNESVPRGRILDRNGKVLV 81  
QY 57 DDAAKALLAYVREKSKAREPDFSNAEKPAD--IQPIDKVTREMAKALVHQFQPHKGYGY 115  
Db 82 DNASKMSITYTRNKTSQKEMINTAKLTLDIKMDTKITER--DK----- 125  
QY 116 FDYKGDINWQWMP-----VKDNEVRWQLHRV--KWWQA 146  
Db 126 ----KDFWIQWYPSAKKLMEKEQLMELEDSISQDQFTQLDRDKIGKKQLTKKDLQV 181



Qy	153	ATGDEKVARWVYQYSVDWARKNPGLSODNDKFVWRPLEVSDRVOSLSPPTPSLSFVN\$PAF	2112
Dd	198	--TSKNLKOWTY-----ASEFGDQGSHGGW-----ECPDFELFPDVGPN	237
Qy	213	TPAFIMEF-----LNSVHQOADLYSTHYAEQG--NHRLF-EAORNLPAGVSYPFEF	259
Dd	238	QKKWMQVSVNGAVSGSGGMQYFVGCDPDGHFKNENPNKVLTWDYGRDFAAVSWSDI	297
Qy	260	--KDSPPR-----WRQTGISVLNTEIKQVYADGMQOFELSIPIHVA	298
Dd	298	PSTD\$RRLLWGMSNWQYANDVPT\$PMWSA--TSIPRELKLKAFTGEGRVVQTPVKELET	355
Qy	299	I-----DIFLKAYGSRAKVNLE-KEPPQS-----YYOTVENMIM--	331
Dd	356	IRGT\$KWKKNLTI\$PASHNVLAQSGSDAYEINAEFKVSPGSAERGFVKRTCGENTKVG	415
Qy	332	----ALISLDPNTMPMG\$SWITDKX---FRMAQFASMARVPFANQAIFYATDGGKQ	383
Dd	416	YDR\$NAKLFDVRSESGNDTFFPAFTGKETAPLKPVNGKVKLRIIVDRSSVFVGNDGKQ	475
Qy	384	GKA---PNFLSKAL---SNAG-----FYFRSGWDKNATVMVLK	416
Dd	476	VITDIILPDR\$SKGLELYANGGVKVKSLTHPLPKVWGTFPFMSNMTCN---TTV----	528
Qy	417	ASPPGEFHAQPDNGTFELFIKGRNFPTDAGVFVYS-----GDEA	455
Dd	529	-----NGTWADTIEBQKQRSGD\$FILSSAGSDTFYESDITIKDGNRGKGAGA	576
Qy	456	IMKL\$RMVROTRIHS\$TLTDNONMVITKARQNKWETGNNLDLVLTNN\$YPNLDHOR\$VL	515
Dd	577	LM-----FR\$KD\$KANGYLANVD\$KHLVKFKFENGGA\$VIAEYKTP-----I	620
Qy	516	FINKXYFL-----VIDRAIGBA-----TG\$NLGVH-WQJ\$ED\$NPVF	550
Dd	621	DVN\$KH\$YLKTEAEGDRFXIYLLDR\$LVIDAHDSV\$FSEGOFGLNVW----DATAVF	670

## RESULT 8

```

US-10-369-493-2315
? Sequence 2315, Application US/10369493
? Publication No. US20030233675A1
? GENERAL INFORMATION:
? APPLICANT: Cao, Yongwei
? APPLICANT: Hinkle, Gregory J.
? APPLICANT: Slater, Steven C.
? APPLICANT: Goldman, Barry S.
? APPLICANT: Chen, Xianfeng
? TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
? TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
? FILE REFERENCE: 38-10(52052)B
? CURRENT APPLICATION NUMBER: US/10/369,493
? PRIOR FILING DATE: 2003-02-28
? PRIOR APPLICATION NUMBER: US 60/360,039
? PRIOR FILING DATE: 2002-02-21
? NUMBER OF SEQ ID NOS: 47374
? SEQ ID NO 2315
? LENGTH: 579
? TYPE: PRT
? ORGANISM: Schizosaccharomyces pombe
US-10-369-493-2315

```

## Query:

QY	171	AKKPLGUSQNDKRFVWKPPLSVSRKUSLPTFLSUFVNSPFTAPFLMEFLNSYHQADY	233
Db	125	SKTNP-----KRDWYFWKPYRNEKGERLPP-----NNMRSYDFT	159
QY	231	LSTHVAGQNHRLFEAQRNLFAGVSYSPFEK--DSPRWOTGISVLNTEIKQVYADGMOFE	289
Db	160	SAWEWDATQ---EYIHLWS-VGQDPLNWETPKVREAVHDIIRFWLDRGV--DG--FR	210
QY	290	LSPIYHVAADIFLKA-----YGSAKRVNLKEPFPQSXYVTVENMINMALI	334
Db	211	LDAI NMISKQRFELADPTTDDRVEYQLAYQVYANGPRIH-----EYUNGIGNI-----	258

QY 335 SISLPDYNPMFGD-SWITDKN-----ERMAQFASWARVFPANQAIKYFATD 380  
 Db 259 ---LTEYDAPSGEMPVLTDTNEILHVGGADRELTWIFQDFVDLDPNQH-KIEGS 314  
 QY 381 GKQKAPNLSK-----ALSNAGF-YTPRSQMDKNATV-MVLKASPPGEPHQAQPDNGTFEL 434  
 Db 315 WELSDLKSLKKWQSALLSGGWNASFIENHDQTRTVSRVLSDSPKYRAYSSKLMALFII 374  
 QY 435 FIKGRNFTPDAGVGVVSGDAIMKLRNWNVYQTRHSTHTLTDNQMVTIKARQNKWETGNN 494  
 Db 375 FQSG---TP-----FVFGQE--LALANIPRDWPIDEYLDVETQNF-----WK----- 412  
 QY 495 LDVLTNPSYPNLDHORSVLFNKYFLVIDRAIGEAATGNLGVHQLKEDSNPVFDKTK 554  
 Db 413 --LFMSGNPQEBI--EKTMDIYNKR-----ARDNGRTPMHW-----DSSPENGFTK 455  
 QY 555 -----NRVYTYRDGNLMIQSLNADRTSLNBEKGKVS-----VYNKELK 595  
 Db 456 AGVKPMWRVTNDYKWN-----AANQVNDPESPYTFWSKALELRKELK 498

RESULT 10

US-10-335-977-8966  
 ; Sequence 8966, Application US/10335977  
 ; Publication No. US20040052799A1

GENERAL INFORMATION:

APPLICANT: DOUGLAS SMITH et al  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES  
 RELATING TO HELICOBACTER PYLORI FOR  
 DIAGNOSTICS AND THERAPEUTICS

NUMBER OF SEQUENCES: 10031

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: LAHIVE & COCKFIELD  
 STREET: 28 State Street

CITY: Boston  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02109-1875

COMPUTER READABLE FORM:  
 MEDIUM TYPE: CD-ROM ISO9660  
 COMPUTER: IBM PC Compatible  
 OPERATING SYSTEM: Windows NT 4.0  
 SOFTWARE: UNIX

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/10/335,977  
 FILING DATE: 30-Dec-2002

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/993,002  
 FILING DATE: 17-DEC-1997

ATTORNEY/AGENT INFORMATION:

NAME: Mandragouras, Amy E.  
 REGISTRATION NUMBER: 36,207  
 REFERENCE/DOCKET NUMBER: GTN-018  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 227-7400  
 TELEFAX: (617) 742-4214

INFORMATION FOR SEQ ID NO: 8966:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 1484 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: YES  
 ORIGINAL SOURCE:  
 ORGANISM: Helicobacter pylori

FEATURE:  
 NAME/KEY: misc\_feature  
 LOCATION: (B) LOCATION 1...1484  
 SEQUENCE DESCRIPTION: SEQ ID NO: 8966:

US-10-335-977-8966

Query Match 3.4%; Score 117.5; DB 12; Length 1484;  
 Best Local Similarity 18.1%; Pred. No. 0.84;

Matches 120; Conservative 100; Mismatches 257; Indels 187; Gaps 29;  
 QY 86 DIRQPIDKVTREMA--DKALVHOFQPHKGYGYDYDKD-INWQWPVKDNEVWQLHRVK 142  
 Db 600 DIVKPSDALKNVDVAALGKQMIGF-----LQODTLNSLESLLQNOQIKSVLDKVL 649  
 QY 143 WQAMALVY-HATGD--EKYAREWVYQYSDWARKNPLGLSQDNDKFWVRPLEVSDRVQSL 199  
 Db 650 AAKGLASIVFQGLDILNPKKKGIFA-----PYGLSQ-----VWQKGFDFNAQG-- 695  
 QY 200 PPTFSLFVNSPAFTPA--FLMEF-----LNSYHQQADYLSHY 235  
 Db 696 ---NVFQNSTESNANGGTLTFNAGNSLIFAGNHNIAFTNHSGLTNLLSNQVSNINVTM 751  
 QY 236 --AEQG-----NHLFEAQRNLFAGVSPFEPKDSRWRQTG----- 269  
 Db 752 LNASGLKINATNNNSVSGNLFINASCVCQSDPTTASATNPFCTTAQNNASSNSNNA 811  
 QY 270 -ISVLNTEIKKQVYADGMQFELSPIYHVAIDIFLKAYGSA----- 309  
 Db 812 PIALNNDESLLVTANGFNS-GNIYANGVVD-FSKIKGSANVKNLVLYNNAQFQANNLT 869  
 QY 310 -----KRVNLEKEFPQSYVQTVF---NMIMAL-ISISLPDYNTPMFGDSW 350  
 Db 870 ISNQAVLEKNASFVTNNLNIQGAFFNNNATOKIEVLQNLVIASNASLSTGYGLEVGG--- 926  
 QY 351 ITDKNFERMAQFASWARVFPANQAIKYFATDGKQKAPNLSKALSNAAGFYTFERSGWDKNA 410  
 Db 927 -ALNNGAIHFNLENSQTPVNPLIQVGGIINLNTQTPFMNVSANGTGT----- 976  
 QY 411 TVMVLKASPPGEPHQAQPDN--GTSELF-----IKGRNFTPDAGVGVVSGDAIMKLRNMYR 464  
 Db 977 ---LLKSSRYIDYNINPNSLSQSYLKLYTLNNGNHNIEKNGVLYLQORVLLQDKGLL 1033  
 QY 465 QTRIHSTLTDNQMVTIKARQNKWETGNNLVDVLTNPSYPNLDHORSVLFNKYFLV 524  
 Db 1034 SVALPNSNNASQNNILSLSVLHNQIKMSYGNKVMDEPTPL-----QDYTVGIQGGALN 1088  
 QY 525 IDRAIGEATGNLGVHW-----QLAKEDSNPVFDKTKNRVYTYRDGNLMIQSLNADRTSLN 580  
 Db 1089 QIEAVG---GNNAIKWLSTLMMETKENPLF-----APIYLENLSLN 1126  
 QY 581 EEEGKVSYYVYNKELKRPAPFVEKPK-KNAGTQNFVSIVYPYDQKAPFIS---IRENKG 636  
 Db 1127 EILG-----VTKDLQNTASLISNPNFNNATSLLEMASYQTQTSRLTKLDFRAREGES 1181  
 QY 637 DFEK 640  
 Db 1182 FSR 1185

RESULT 11

US-10-282-122A-59042  
 ; Sequence 59042, Application US/10282122A  
 ; Publication No. US20040029129A1

GENERAL INFORMATION:

APPLICANT: Wang, Liangsu  
 APPLICANT: Zamudio, Carlos  
 APPLICANT: Malone, Cheryl  
 APPLICANT: Haselbeck, Robert  
 APPLICANT: Ohlsen, Karl  
 APPLICANT: Zyskind, Judith  
 APPLICANT: Wall, Daniel  
 APPLICANT: Trawick, John  
 APPLICANT: Carr, Grant  
 APPLICANT: Yamamoto, Robert  
 APPLICANT: Forsyth, R.  
 APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
 FILE REFERENCE: ELITRA.034A  
 CURRENT APPLICATION NUMBER: US/10/282,122A  
 CURRENT FILING DATE: 2003-02-20  
 PRIOR APPLICATION NUMBER: 60/191,078



```
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59042
; LENGTH: 2902
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-10-282-122A-59042

Query Match          3.4%; Score 117.5; DB 12; Length 2902;
Best Local Similarity 18.1%; Pred. No. 2.4;
Matches 120; Conservative 100; Mismatches 257; Indels 187; Gaps 29;

Qy 86 DIRQPIDKVTREMA--DKALVHQFPHKGYGYFDYKGD-INMQMVPKDNVNRWQLHRVK 142
Db 2018 DIVKPSDALKNDVAALGKQMGIEF-----LGQDTLSLESLLQNOQIKSVLDKVL 2067

Qy 143 WQAMALVY-HATGP--EKYAREWVQYSDWARKNPLGLSQDNDKFVWRPLEVSDRVQSL 199
Db 2068 AAKGIGSIYEQGLGDLIPNLGKKGIFA-----PYGLSQ-----VMQKGDFFSNAQG-- 2113

Qy 200 PPTFSLFVNSPAFTPA--FLMEF-----LNSVHQOQADYLSTHY 235
Db 2114 ----NVFQNSFTSNANGGTLTSFNAGNSLIPAGNHHIAFTNHSGLTLLSNQVSNINVTM 2169

Qy 236 --AEGG-----NHLRFAQRNLPAQVSPPEPKDSRWRQTG----- 269
Db 2170 LNASGLKINATNNVSVSQGLFNASCVOQSDPTTASATNPCTTAQNNASSNASNNA 2229

Qy 270 -ISVLTNTEIKQVYADGMQFELSPIYHVAADIFLKAYGSA----- 309
Db 2230 PLALNNDESLVVTANGFNFS--GNLYANGVVD-FSKIKGSANVKNLYLYNNAQFOANNLT 2287

Qy 310 -----KRYNLEKEFPQSYVQTVF-----NMIMAL--ISISLPDYNTPMFGDSW 350
Db 2288 ISNQAVLEKNASFVTTNLIQGAFFNNAQTKEVLQNLVIASNASLSGTGIYGVGG--- 2344

Qy 351 ITDKNFRMAQASWARVPANQAIVKFDGQKQKAPNPLSKALSNAGFYTPRSQWDKNA 410
Db 2345 -ALANNLGAHFLENSTQPVNPLIQVGGIINLTNTTQTPFMNVSVANGGTYT----- 2394

Qy 411 TVMVLKASPPGEFHAQPN--GTFELF-----IKGNFTPDAGVFYVSGDEALMKLRNWR 464
Db 2395 ---LKSRYIDYNINPNSLQYLKLTLYLNINGNHIIEKNGVLVYLGQRLVLLQDKGLL 2451

Qy 465 QTRIHSTLTLDQNQNVITKARQNKWETGNNLDVLTATYTPSPYNPLDHRQSVLPFIKKYFLV 524
Db 2452 SVALPNSNNAQNNILSLSLVHLNQIKMSYGNKVMDFTPPLL-----QDYIVIGIQCSALN 2506

Qy 525 IDRAITGEATNGLGVW-----QKEDSNVPVDTKRVYTYTDRDGNMLIQSLNADRTSLN 580
Db 2507 QIEAVG---GNNAIKWLSTLMMETKENPLF-----APIYLENHSIN 2544

Qy 581 BEEGKSVYVYNKELKAPAFVEKPK-KNAGTQNFVSIVVPYDGKAPFIS---IRENKN 636
```

```
Db 2545 EILG-----VTKDLQNTASLISNPNFRNATSLLEMASYTQOTSRLLTKLSDFRAREGSN 2599
Qy 637 DPEK 640
Db 2600 FSR 2603

RESULT 12
US-10-452-024-107
; Sequence 107, Application US/10452024
; Publication No. US20040013687A1
; GENERAL INFORMATION:
; APPLICANT: Simpson, Lance
; APPLICANT: Park, Jung-Beak
; APPLICANT: Maksymowich, Andrew
; TITLE OF INVENTION: Compositions and Methods For Trans epithelial Molecular Transport
; FILE REFERENCE: 9855-96U1
; CURRENT APPLICATION NUMBER: US/10/452,024
; CURRENT FILING DATE: 2003-06-02
; PRIOR APPLICATION NUMBER: 60/384,949
; PRIOR FILING DATE: 2002-05-31
; NUMBER OF SEQ ID NOS: 188
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 107
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-10-452-024-107

Query Match          3.3%; Score 117; DB 15; Length 1163;
Best Local Similarity 17.1%; Pred. No. 0.64;
Matches 141; Conservative 129; Mismatches 280; Indels 274; Gaps 38;

Qy 17 LSSGNILAQSSSIIRKDF-----DHNLEYSGLKVKVKAAGNYDAAKALIA 65
Db 380 LKMSNVYDGLKGTWNNFYAVYKIPYNGIDVHINYSYLNVN--VEEINNIPPINDADIY 438

Qy 66 YREKSKAREPDFS-----NAEKPADIROPIDKVT-----EMADKALV 104
Db 439 PYRKNSDPFIPYINITETKEINTTTPLSVNYLQAQVTSNDINLSLDSKVISLKDRLV 498

Qy 105 HQFQHKGYGYFD--YKGDIN-----WQMPVKDNVNRWQLHRVKW 144
Db 499 YSFLDNT-IDYLDSTIKYDEPINTDKYYLWLKELFPRYSFDMTEIQEWNIPCGINKVVPW 557

Qy 145 QAMALVYHATGDEKYAREWVQYSDWARKNPLGLSQDNDKFVWRPLEVSDRVQSL----- 199
Db 558 LGKALNILNTGNS-----FIEFKTLGPISLINKKENIIMPKIEIDEIPNSMLNSF 609

Qy 200 ----PPTFSLFVNSPAFTPAFLMEFINSYHQOQADYLSTHY-----AEQ----- 238
Db 610 KDLSENLENIFSKNSYFEKIYYDFLDQWMTQ--YYSQYFDLICMAKRSVLQAESLIKKI 667

Qy 239 -----GNHRLFE-----AQRN-----LFAQVS 255
Db 668 IQKLSYILGNSISADNLVLMNLTTTLTLDISNESQIAMNVDSFLNSAAICVFEGNI 727

Qy 256 PPEF-----KDSRWRQTGISVLNTE-----IKQVYADGMQFELSPIYHVA 299
Db 728 YPKFISFMEQCINNKNKTREFIQKCTNITENEKQLINRNIFS-SLDFDFLINIEN---- 782

Qy 300 DIFLKAYGSAKRVNLEKEFPQSYVQTVENMMALISISLPDYNTPMFGDSWITDKNFRMA 359
Db 783 ---LKSLSFSSTALLIKEETSPYE-----LVLYAQFQEPDNN--IGDA--SAKNTSIE 828

Qy 360 QFASWARYF-----PANQAIKYFATDGKQKAPNPLSKALSNAGFYTPRSQWDKN 409
Db 829 YSKDIDLIVGINDALYNGANQSISF-----SNDFFENGLTNSFSYIF---WLRN 876

Qy 410 ATVMVLKASPPGEFHAQPDNGTFFELFKGRNFTPDAGVVF---YSGDEALMKLRNWRQ 465
Db 877 LGKDTIKSLIG--SKEDNCGWEIYFQ-----DTGLFVNMDISNGNEKNIYLSDSVNN 927
```



```
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-10-282-122A-70737

Query Match      3.3%; Score 116; DB 12; Length 919;
Best Local Similarity 16.2%; Pred. No. 0.55;
Matches 124; Conservative 84; Mismatches 216; Indels 258; Gaps 31;

QY 56 YDDAAKALLAYREKSKAREPDFSNAEKPADIR-----QPIDKVTREMADK 101
DB 370 YDKANKAVTDLLEKADLLKDFITHSYPHDWTCKPVIFRATPOWFAKINKVRQDILDA 429
QY 102 ALVHQFQPHKGYPDYKDNWQMPVKDNEVRWQLHRVQWQAMALVYHA-TGDEKXA 160
DB 430 IEDTNFK-----VDMGKTRIYNN--IRDRG-EWVISRQVRWGVPLPVFAENGDIIMT 479
QY 161 REWVYQYSDWARKNPLGLSQNDKEVWRPLEVSDRVQSLPPTFSLFVNSPAPTFAFLMEF 220
DB 480 KETVNHVAD-----LFEKHGNSIWFEKEAK---ELLPEGFS-HPGSPN----- 518
QY 221 LNSVHQQADYLSTHYAEOGNHR-LPEAQRNLFAGVSFPE--FKDSPRWRTGISVLNTE 276
DB 519 -GEFTKETDIMDVWDSSSHRGVLETRPEL-----SFPADLYFEGSDQYR---GWENSS 569
QY 277 IKQVYADGMOPELSPIYHVAIDIFLKAYGSAKRVNLEKEPPQSYQVOTVENMINALISI 336
DB 570 ITTAVATRGQ-----APYKFLLSHG----- 589
QY 337 SLPDYNTPMFGDSWITDKNFRMAQFASWARVPFANQAIKYPATDCKQKQKAPNFLSKALSN 396
DB 590 -----FVMDG-EGKK-----MKSLSLN 605
QY 397 AGFYTFRSQWKNATVMVLKASPPGFEHAQPDNGTFELFKGRNFTPDAGVFVYSGDEAI 456
DB 606 -----VIV-----PDQVVKQKGADIAVLWVSSTDVLSVRI-----SDEIL 641
QY 457 MKLRNWRQTR-----IHSTLTLDNQNNVITKARQNKWETGNNLDV 497
DB 642 KQTSVYRKIRNTLRFMLGNINDFNPETDSTAETNLLVEDRYLLNRLREFTASTINNYEN 701
QY 498 LTYTN--PSYPN-----LDHORSVLFINKK-----YFLVIDRAIGBAT 533
DB 702 FDYLNIEVQNNFINVELSNFYLDYKGKDLILYIEKKDSHKRSMQVVLQOILVD--MTKLL 759
QY 534 GNLGVH-----W-----OLKEDSNPVFDKTKN-----RVYTTYRDGNNLMIOQLN 573
DB 760 APILVHTAEVWMSHTPHVKEESVHLSMDPKVVDVDEELLEKWNTFMNLRDDVNRALEQAR 819
QY 574 ADRTSLNEEGKVSIVYNKELKRPAPVFE-----KPKKNAGTQNFVSIVY 618
DB 820 NEKVIGKSLEAKVWIGSNESFNFAEFLOQFNDLQOLFIVSQVEVKQKNDG-----VSY 873
QY 619 PYDGQKAPESIRENKGNDFEK 640
DB 874 QYG-----DIHLKHAEGEKCEK 890

RESULT 15
US-09-802-640-32
; Sequence 32, Application US/09802640
; Publication No. US20030036057A1
; GENERAL INFORMATION:
; APPLICANT: Braun, Andreas
; APPLICANT: Bonsal Aruna
; APPLICANT: Klevn Patrick
; TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND THEIR USE
; FILE REFERENCE: 24736-2048
; CURRENT APPLICATION NUMBER: US/09/802,640
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
```

```
; LENGTH: 4563
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-802-640-32

Query Match      3.3%; Score 116; DB 10; Length 4563;
Best Local Similarity 16.3%; Pred. No. 6.7; Indels 244; Gaps 26;
Matches 104; Conservative 94; Mismatches 94;

QY 218 MEFINSY-----HQQADY-----LSTHYAEOGNHRLFEA-----QRNL----- 250
DB 3068 IDFLNNYALFLSPSAQQAQASQVSAFFNQYKYNQNFAGNENIMEAHVINGEANLDELN 3127
QY 251 -----FAGVSFPEFKDSPRWRTG----- 269
DB 3128 IPLTIPEMRPLFYTIITTPDLKDFSLWEKTKFLKTKTKQSFDSLVSVAQYKKNKHSIT 3187
QY 270 --ISVLNTEIKQVYADGMOPELSPIYHVAIDIFLKAYGSAK---RVNLEK---RFPQ 320
DB 3188 NPLAVLCEFIQSQIKSFRHFEKN---RNNALDFVTKSYNETKIKFKYKAEKSHDELPR 3244
QY 321 SYVQTVENMINALISISLDPDYNTPMFGDSWITDKNFRMAQF-----ASWARVPFANQAIKY 376
DB 3245 TF-----QIPGYTPVW-----NVEVSPFTIEMSAFGYVFPKAVSMES 3282
QY 377 FATDGKQKAPNF-----LSKALSNAGF-YT 401
DB 3283 FSIILGSDVRVPSYTIILPSELPLVHLVPRNLKLSLPHFKELCTISHIFIPAMGNITYDFS 3342
QY 402 FRSGWDKNATVMVLKASPPGFEHAQPDNGTFELFKGRNFTPDAGVFVYSGDEAI----- 456
DB 3343 FKSS-----VITLNTNAELFNQSD--IVAHLSSSSSVIDALQYKLEGTRTRTKRG 3392
QY 457 -----MKLRNWRQTRIHSTLTLDNQNNVITKARQNKWE-----TGNLNDVL 498
DB 3393 LKLATALSLNKFVEGSHNSTVSLTTKNWEYSVAKTTKAEIPLRMNFKQELNGTKSKP 3452
QY 499 TYTNPSYNLDHORSVLFINKK-----YFLVIDRAIGBATGNL----- 536
DB 3453 TVSSSMFEKYPDNFSMLYSTAKGAVDVKHLSLESLSITSFIESSTKGDYKSVLSREYSGT 3512
QY 537 -----GVH-----WQIKEDSNPVFDKTKNVRVYTYRDG--NNLM 568
DB 3513 TASEANTVINKSTRSSVKLOQTSKIDDIWNLEVENFAGEATLQRIISLWHEHSTXNHLQ 3572
QY 569 IQSL---NADRTS---LNEEEGKVSIVYNKELKRPAPVFEKP-----KKNAGTQN----- 612
DB 3573 LEGLFTTNGEHTSKATLELSPWQMSALVQVHASQPSSEFHPDPLDGLQEVALNANTKNQIR 3632
QY 613 FVSIVYPYDGQKAPESIRENKGNDFEKGNLTLTING 651
DB 3633 WKNEVRIHSGSFQSQVEL-----SNDQEKAHLDIAGSLEG 3667

Search completed: July 27, 2004, 17:33:03
Job time : 53 secs
```

This Page Blank (uspto)